

DELPHION

No active tr.

Select CR**RESEARCH****PRODUCTS****INSIDE DELPHION****Log Out** **Work Files** **Saved Searches**

My Account

Search: Quick/Number Boolean Advanced Der

The Delphion Integrated ViewGet Now: ☒ PDF | [File History](#) | [Other choices](#)Tools: Add to Work File: [Create new Work File](#)View: [INPADOC](#) | Jump to: [Top](#)  Go to: [Derwent](#) [Ema](#)Title: **JP09083655A2: VOICE INTERACTIVE SYSTEM**Derwent Title: Audio interaction system used in computer network - incorporates server for controlling on-line interactive communication between different terminals connected to communication circuit [\[Derwent Record\]](#)

Country: JP Japan

Kind: A

Inventor: KAWAGUCHI NAOHISA;
MATSUI KAZUKI;
ONO TAKASHI;
IWAKAWA AKINORI;
HARADA HIROAKI;Assignee: FUJITSU LTD
[News, Profiles, Stocks and More about this company](#)

Published / Filed: 1997-03-28 / 1995-09-14

Application Number: JP1995000237079

IPC Code: IPC-7: [H04M 11/00](#); [H04M 3/56](#);

Priority Number: 1995-09-14 JP1995000237079

Abstract: PROBLEM TO BE SOLVED: To appropriately control inter-user communication quantity by providing a memory buffering a voice signal which is transmitted from a terminal by means of a server and a modulator controlling a voice signal distribution which is executed buffering.

SOLUTION: A proposal transmitted from the respective terminals is stored in buffer memories 21A and 21B inside the server 20 through the use of a control part 23, the respective proposals are properly given order by a scheduler 22 and distributed to the respective terminals A and B in order. Thus, a problem that the proposal is unnaturally interrupted by the parallel incoming of the proposals to the respective terminals A and B when the proposals are simultaneously executed from the plural terminals A and B so as to exceed the band width of a line 1 can be evaded. The handling time of a processing such as the parallel incoming of the plural proposals or the plural proposals are mixed in the respective terminals A and B is saved.

COPYRIGHT: (C)1997,JPO

INPADOC Legal Status: None Get Now: [Family Legal Status Report](#)Family: [Show 2 known family members](#)Other Abstract: [DERABS G97-252129](#) [DERG97-252129](#)

Info:



[Nominate this for the Gallery...](#)



Copyright © 1997-2006 The Tho

[Subscriptions](#) | [Web Seminars](#) | [Privacy](#) | [Terms & Conditions](#) | [Site Map](#) | [Contact U](#)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-83655

(43) 公開日 平成9年(1997)3月28日

(51) Int.Cl.⁶

H 0 4 M 11/00
3/56

識別記号

庁内整理番号

9465-5G

F I

H 0 4 M 11/00
3/56

技術表示箇所

Z

審査請求 未請求 請求項の数25 O L (全 62 頁)

(21) 出願番号

特願平7-237079

(22) 出願日

平成7年(1995)9月14日

(71) 出願人

000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者

川口 尚久

神奈川県川崎市中原区上小田中1015番地
富士通株式会社内

(72) 発明者

松井 一樹

神奈川県川崎市中原区上小田中1015番地
富士通株式会社内

(74) 代理人

弁理士 山田 正紀

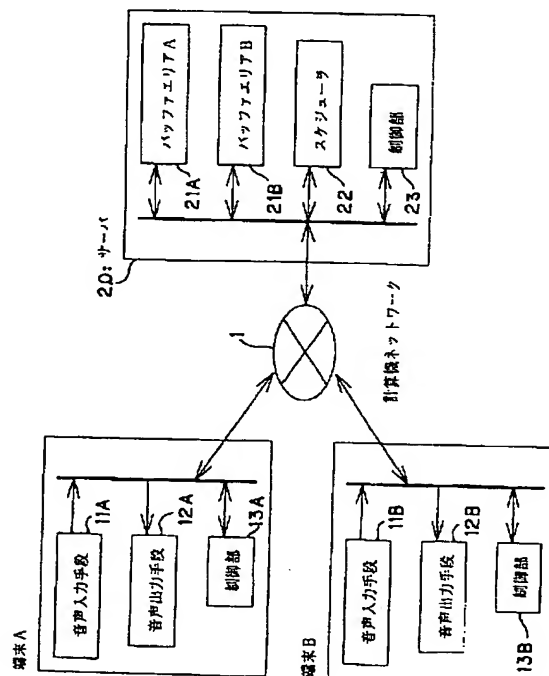
最終頁に続く

(54) 【発明の名称】 音声対話システム

(57) 【要約】

【課題】 本発明は、コンピュータおよびコンピュータネットワークを利用したオンラインチャットシステムにおける、音声によって複数のユーザが対話する音声対話システムに関し、ユーザ間の通信量を適切に制御する。

【解決手段】 音声を音声信号に変換する音声入力手段および音声信号を音声に変換する音声出力手段を備えた、通信回線に接続される複数の端末と、通信回線を介してこれら複数の端末と接続され、該端末との間で音声信号の収集、配信を行なうサーバとを備え、前記サーバが、前記端末から送信されてきた音声信号をバッファリングするメモリと、該メモリにバッファリングされた音声信号の配信を制御するスケジューラとを備えた。



【特許請求の範囲】

【請求項1】 音声を音声信号に変換する音声入力手段および音声信号を音声に変換する音声出力手段を備えた、通信回線に接続される複数の端末と、通信回線を介してこれら複数の端末と接続され、該端末との間で音声信号の収集および配信を行なうサーバとを備え、これら複数の端末間で音声を用いて対話を行なう音声対話システムにおいて、

前記サーバが、前記端末から送信されてきた音声信号をバッファリングするメモリと、該メモリにバッファリングされた音声信号の配信を制御するスケジューラとを備えたことを特徴とする音声対話システム。

【請求項2】 前記スケジューラが前記メモリにバッファリングされた各音声信号の配信順序を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項3】 前記サーバが、音声を同時に送信する同時発言者数管理手段を備え、前記端末からの、所定の数を超える音声信号の同時送信を禁止するものであることを特徴とする請求項1記載の音声対話システム。

【請求項4】 前記サーバもしくは前記端末が、音声信号の無音部分を検出する無音部分検出手段を備え、前記スケジューラが、音声信号が無音部分により分離されてなる各発言単位毎に配信順序を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項5】 前記スケジューラが、音声信号が前記サーバに到着した時刻順に該音声信号が該サーバから配信されるように、音声信号の配信順序を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項6】 前記サーバが、前記端末の内部時刻の送信を指示するコマンドを該端末に送信する手段を備え、前記端末が、前記コマンドに応答して自分の内部時刻を前記サーバに送信する手段と、前記音声信号を前記サーバに送信するに際して、該音声信号に送信時の自分の内部時刻を付して送信する手段とを備え、前記スケジューラが、音声信号が前記端末から発信された時刻順に該音声信号が前記サーバから配信されるように、音声信号の配信順序を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項7】 前記サーバが音声信号をミキシングするミキサを備えたことを特徴とする請求項1記載の音声対話システム。

【請求項8】 前記スケジューラが、音声信号が前記サーバに到達した時刻に基づいて、音声信号の配信時刻を制御するものであることを特徴とする請求項7記載の音声対話システム。

【請求項9】 前記サーバが、前記端末の内部時刻の送信を指示するコマンドを該端末に送信する手段を備え、前記端末が、前記コマンドに応答して自分の内部時刻を

前記サーバに送信する手段と、前記音声信号を前記サーバに送信するに際して、該音声信号に送信時の自分の内部時刻を付して送信する手段とを備え、

前記スケジューラが、音声信号が前記端末から発信された時刻に基づいて、音声信号の配信時刻を制御するものであることを特徴とする請求項7記載の音声対話システム。

【請求項10】 前記端末それぞれが論理上の人物像を有し、前記スケジューラが、前記論理上の人物像空間内における該人物像の論理上の位置関係、距離、周囲環境、視線方向、視界、姿勢のうちの少なくとも1つに基づいて音声信号の配信を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項11】 前記端末それぞれが論理上の人物像を有し、前記論理上の人物像空間内における、該人物像の位置関係、距離、周囲環境、視線方向、視界、姿勢のうちの少なくとも1つに基づいて音声信号に変更を加えるエフェクタを備えたことを特徴とする請求項1記載の音声対話システム。

【請求項12】 前記サーバが、前記複数の端末のうちの対話を行なう対話グループに属する端末を登録する1以上のグループ管理テーブルを備え、

前記スケジューラが、受信した音声信号が、該音声信号を発信した端末と同一の対話グループに属する端末に配信されるように、音声信号の配信を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項13】 前記端末が、自分が送信する音声信号を受信する受信先の端末を指定する複数の発信モードのうちのいずれか1つの発信モードを選択する発信モード選択手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項14】 前記端末が、自分が受信する音声信号の送信元の端末を指定する複数の聴取モードのうちのいずれか1つの聴取モードを選択する聴取モード選択手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項15】 前記端末が、前記サーバに向けて、音声信号の発信元である端末のうちの、自分あてに送信される音声信号の受信を拒否する端末を指定する割込禁止通知を行なう割込禁止通知手段を備え、

前記サーバが、割込禁止通知により指定された送信元の端末から受信した音声信号の送信先が該割込禁止通知を行なった端末であった場合に、該送信先の端末への該音声信号の転送を取り止めるものであることを特徴とする請求項1記載の音声対話システム。

【請求項16】 前記対話グループに対応してパスワードが付されており、前記端末が、パスワードを入力して前記サーバに送るパスワード入力手段を備え、前記サーバが、端末から送られてきたパスワードを検査して正しいパスワードを送信してきた端末を、該パスワ

ードに対応する対話グループに加入させるものであることを特徴とする請求項1記載の音声対話システム。

【請求項17】 前記端末が、話者と、聴取者を表示する表示手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項18】 前記サーバが、音声信号の音質を変化させるエフェクタを備えたことを特徴とする請求項1記載の音声対話システム。

【請求項19】 前記サーバが、所定の効果音を表わす効果音信号を格納する効果音バッファを備え、前記スケジューラが、送信されてきた音声信号の配信を制御するとともに、前記効果音バッファに格納された効果音信号の配信を制御するものであることを特徴とする請求項1記載の音声対話システム。

【請求項20】 前記サーバもしくは前記端末が、音声信号をテキストデータに変換する音声認識手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項21】 前記サーバと前記端末がアナログ回線で接続され、アナログ信号としての音声信号の始端および終端に音声信号の始端、終端を識別するためのトーン音を発生させるとともに受信したトーン音を検出する、それぞれトーン音発生手段およびトーン音検出手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項22】 前記サーバが、前記端末どうしの対話を記憶する記憶手段を備えたことを特徴とする請求項1記載の音声対話システム。

【請求項23】 前記サーバが、前記端末どうしの対話を前記記憶手段に記憶するにあたり、該対話に付随する付属情報を付加して記憶するものであることを特徴とする請求項22記載の音声対話システム。

【請求項24】 前記サーバが、前記付属情報に基づいて前記記憶手段に記憶された対話中の、所望の対話部分を検索する検索手段を備えたことを特徴とする請求項23記載の音声対話システム。

【請求項25】 対話の話速を変換する話速変換手段を備えたことを特徴とする請求項22記載の音声対話システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、コンピュータおよびコンピュータネットワークを利用したオンラインチャットシステムにおける、音声によって複数のユーザが対話する音声対話システムに関する。

【0002】

【従来の技術】コンピュータネットワークの普及により、ネットワーク上で複数のユーザが会話することで、作業を進めたり、ゲームを行ったり、生活情報を交換するということが容易になってきた。このため、一般ユーザでもこのような、ネットワーク会話システム（チャ

ットシステム）に参加できるように、キーボードに換わる、もっと簡単な入力手段が求められている。

【0003】従来のチャットシステムにおいては、キーボードを操作することで、文章を入力し、これを送信することで会話を行っていた。ところが、一般ユーザにとっては、かな漢字変換などのキーボード操作は非常に難しいことであり、このため入力速度が遅くなり、会話の流れについていけないという問題点があった。この問題を解決するために、音声により会話するシステムも開発されている。

【0004】図61は、従来の音声対話システムの構成図である。ここでは簡単のため、端末は端末Aと端末Bとの2台のみ示されており、これらの端末A、B間で対話が行なわれるものとする。各端末A、Bには、音声音声信号に変換する、マイクロホン等の音声入力手段11A、11B、および音声信号を音声に変換する、スピーカ等の音声出力手段12A、12Bが備えられており、通信回線1を介して、ルータ2の制御に基づいて端末Aと端末Bが接続され、音声信号を送受信することにより端末A、B間での対話が行なわれる。

【0005】しかしながら、従来のこのような音声対話システムでは、通信路のバンド幅が不足することから、2者～4者間で対話を行なうことのできるシステムしか存在しない。

【0006】

【発明が解決しようとする課題】従って、複数のユーザがネットワークを通して会話するためには、ユーザの増加に従って増える音声データをどのようにして削減するかが課題である。本発明は、上記事情に鑑み、ユーザ間の通信量を適切に制御するのに好適な音声対話システムを提供することを目的とする。

【0007】

【課題を解決するための手段】上記目的を達成する本発明の音声対話システムは、音声音声信号に変換する音声入力手段および音声信号を音声に変換する音声出力手段を備えた、通信回線に接続される複数の端末と、通信回線を介してこれら複数の端末と接続され、これらの端末との間で音声信号の収集および配信を行なうサーバとを備え、これら複数の端末間で音声を用いて対話を行なう音声対話システムにおいて、上記サーバが、端末から送信されてきた音声信号をバッファリングするメモリと、そのメモリにバッファリングされた音声信号の配信を制御するスケジューラとを備えたことを特徴とする。

【0008】ここで、上記スケジューラは、音声信号の配信順序を制御するものであってもよい。図1は、この態様の原理図である。この図、および、一部の図を除き以下の各図においても、端末は代表的に2台のみ図示する。

【0009】ここでは、図1に示すように、それぞれの端末A、Bより送信されて来た発言を、制御部23を用

いて、サーバ20内のバッファメモリ21A、21Bに記憶し、それぞれの発言をスケジューラ22により適当に順序付けし、順に、各端末A、Bに配信する。ここでは、解りやすさのため、バッファメモリは、各端末A、Bに対応させて端末数と同数図示してある。

【0010】これにより、複数の端末A、Bより同時に発言が行われた場合に、各端末A、Bにそれらの発言が並行して着信することで回線1のバンド幅を越えてしまい、発言が不自然に途切れたりするという問題点を回避することができる。また、各端末A、Bにおいて複数の発言が平行して着信する、あるいは各端末A、Bにおいて複数の発言をミキシングするという処理の手間も省くことができる。尚、この図1には、各端末A、Bを制御する制御部13A、13Bも図示されている。

【0011】本発明の音声対話システムにおいて、上記サーバが、音声と同時に送信する同時発言者数管理手段を備え、上記端末からの、所定の数を超える音声信号の同時送信を禁止するものであることが好ましい。図2は、この態様を示す原理図である。ここには、図2に示すように、それぞれの端末A、Bより発言を行う場合、まず、発言要求手段14A、14Bにより、サーバ20に対して発言したいという要求を伝える。サーバ20では、同時発言者数管理手段24に、予め、同時に発言できる人数を回線1のバンド幅などを考慮して設定しておく。例えば、3人までなら、同時発言者数管理手段24に3を設定しておく。端末より発言要求が来たならば、この同時発言者数管理手段24に保持されている数値を確認し、1以上なら、その端末A、Bに許可通知を与えるとともに、数値を-1する。もし、数値が既に0となっていたならば、どれかの端末から発言終了通知が来るまで待つ。

【0012】端末側では、許可通知を受信したら、発言許可表示手段15A、15Bにより、ユーザに発言を行っても良いことを知らせる。ユーザはこの表示を確認の後、発言を開始し、発言完了後に、発言終了通知手段16A、16Bにより、サーバ20に対し発言終了を通知する。サーバ側では、発言終了通知を受信したならば、同時発言者管理手段24に保持されている数値を+1する。

【0013】本発明の音声対話システムでは、サーバもしくは端末が、端末から送信されてきた音声信号の無音部分を検出する無音部分検出手段を備え、スケジューラが、音声信号が無音部分により分離されてなる各発言単位毎に配信順序を制御するものであることが好ましい。図3は、この態様を示す原理図である。

【0014】図3に示すように、それぞれの端末より送信されてきた発言を、制御部23を用いて、サーバ20内のバッファメモリ21A、21Bに記憶し、無音部検出手段に25により、バッファメモリ21A、21B内のPCMデータを調べることで、音量の大きさを計

測し、音量が予め設定された値より小さい期間が、予め設定された時間より長ければ、この部分で発言を分割し、それぞれの発言をスケジューラにより適当に順序付けし、順に、すべての端末に配信する。

【0015】これにより、本発明の基本的な利点に加え、発言の切れ目を見付けて、その切れ目で他の発言に切り替えることができるので、無音部分で無駄に占有されることを防ぐことが可能となる。本発明の音声対話システムでは、上記スケジューラが、音声信号がサーバに到着した時刻順に音声信号がサーバから配信されるように、音声信号の配信順序を制御するものであることも好ましい態様である。

【0016】図4は、この態様を示す原理図である。図4に示すように、それぞれの端末より送信されてきた発言を、制御部23を用いてサーバ20内のバッファメモリ21A、21Bに記憶し、さらに、発言の受信開始時の時刻を、各バッファメモリ21A、21Bに記憶しておく。スケジューラ22においては、もし、現在、送信中の発言があれば、それを続行する。もし、送信完了していれば、全てのバッファメモリ21A、21Bの中で、受信開始時刻の早いものを、発言の区切りまで、すべての端末に送信する。

【0017】これにより、本発明の基本的な利点に加え、さらに、時間軸に沿った会話が可能になるという利点がある。本発明の音声対話システムでは、上記サーバが、端末の内部時刻の送信を指示するコマンドを端末に送信する手段を備え、端末が、そのコマンドに応答して自分の内部時刻をサーバに送信する手段と、音声信号をサーバに送信するに際して、その音声信号に送信時の自分の内部時刻を付して送信する手段とを備え、上記スケジューラが、音声信号が端末から発信された時刻順に音声信号がサーバから配信されるように、音声信号の配信順序を制御するものであることも好ましい態様である。

【0018】図5は、この態様を示す原理図である。図5に示すように、端末A、Bがサーバ20に接続された時点で、サーバ20内の内部時間補正值計算部27が起動され、まず、サーバ20より、端末A、Bに向かって、端末A、Bのローカルな時刻を内部時計17A、17Bより読み取り、その時刻をサーバ20に対して送信するよう指示する。この指示を出した時刻を T_s とする。サーバ20では、この指示に応答して、端末A、Bのローカルな時刻が端末A、Bより報告されるのを待つ。端末A、Bから報告が得られたら、この時刻を T_e とし、報告された端末A、Bのローカルな時刻を T_1 とすると、サーバ20の内部時計26とある端末A、Bの内部時計17A、17Bの差分は、内部時間補正值 $T_d = T_s + (T_e - T_s) / 2 - T_1$ となる。この値をすべての端末A、Bについて計算して、各バッファ21A、21Bの内部時間補正值メモリに記録する。

【0019】各端末A、Bにおいて、発言が行なわれた

時刻 T_x を内部時計17A, 17Bにより計測し、その時刻を付加してサーバ20に送信する。これにより、サーバ20において、各端末A, Bから送信されてきた発言の実際に行われた時刻を、 $T_x - T_d$ により、サーバ内部時計の時刻に補正することができる。各端末A, Bからの発言をバッファ21A, 21Bに格納するときには、端末A, Bより送信された内部時刻を、内部時間補正值を用いて、サーバ内部時刻に補正し、補正済受信開始時刻メモリに記憶する。スケジューラ22は、発言が実際に行われた順に従って、各発言を順次、すべての端末A, Bに送信することができる。

【0020】これにより、本発明の基本的な利点に加え、より厳密に誰が最初に発言を行なったかを決定することができるので、早いもの順で行なうサービス、例えば、クイズのようなサービスも公正に行なうことが可能になる。本発明の音声対話システムでは、上記サーバが音声信号をミキシングするミキサを備えることが好ましい。

【0021】図6は、この態様を示す原理図である。図6に示すように、それぞれの端末A, Bより送信されてきた発言を、サーバ20内のバッファメモリ21A, 21Bに記憶し、それぞれの発言をミキサ29により、1つの音声にミキシングして一旦出力バッファエリア28に格納した後、すべての端末A, Bに配信する。

【0022】これにより、複数の端末A, Bより同時に発言が行なわれた場合に、各端末A, Bに、それらの発言が並行して着信することで、回線1のバンド幅を越えてしまい、発言が不自然に途切れたりするという問題を回避することができる。また、各端末A, Bにおいて、並行して着信する、あるいは、複数の発言をミキシングするという処理の手間も省くことができる。

【0023】本発明の音声対話システムでは、上記スケジューラが、音声信号がサーバに到達した時刻に基づいて、音声信号の配信時刻を制御するものであることも好ましい態様である。図7は、この態様を示す原理図である。図7に示すように、それぞれの端末A, Bより送信されてきた発言を、サーバ20内のバッファメモリ21A, 21Bに記憶し、さらに発言の受信開始時の時刻を、各バッファ21A, 21Bに記憶しておく。ミキサ29は、受信開始時刻を参照して各発言を時系列に揃え、同時刻の音をミキシングして出力バッファエリア28に一旦格納する。スケジューラ22は、同時刻の発言はミキシングされて、各発言を時刻順に配信する。

【0024】これにより、本発明の基本的な利点に加え、各発言を時刻順に揃えてミキシングすることが可能になる。本発明の音声対話システムでは、上記サーバが、端末の内部時刻の送信を指示するコマンドを端末に送信する手段を備え、端末が、そのコマンドにตอบสนองして自分の内部時刻をサーバに送信する手段と、音声信号をサーバに送信するに際して、音声信号に送信時の自分の

内部時刻を付して送信する手段とを備え、上記スケジューラが、音声信号が端末から発信された時刻に基づいて、音声信号の配信時刻を制御するものであることも好ましい態様である。

【0025】図8は、この態様を示す原理図である。図8に示すように、端末A, Bがサーバ20に接続された時点で、まず、サーバ20より、端末A, Bに向かって、端末A, Bのローカルな時刻を内部時計17A, 17Bより読み取り、その時刻をサーバ20に対して送信するように指示する。この指示を出した時刻を T_s とする。サーバ20では、この指示にตอบสนองして、端末A, Bのローカルな時刻が端末A, Bより報告されるのを待つ。端末A, Bから報告が得られたら、この時刻を T_e とし、報告された端末A, Bのローカルな時刻を T_1 とすると、サーバ20の内部時計26とある端末A, Bの内部時計の差分は、 $T_d = T_s + (T_e - T_s) / 2 - T_1$ となる。この初期設定処理を各端末A, Bについてまず行なう。

【0026】各端末A, Bにおいて、発言が行なわれた時刻 T_x を内部時計17A, 17Bにより計測し、その時刻を付加してサーバ20に送信する。これにより、サーバ20において、各端末A, Bから送信されてきた発言の実際に行なわれた時刻を、 $T_x - T_d$ により、サーバ内部時計の時刻に補正することができる。このようにすることで、各端末A, Bからの発言をバッファ21A, 21Bに格納するとき、サーバ内部時計に補正した発言開始時刻を付加して記憶する。これにより、ミキサ29は、発言が実際に行なわれた時刻に従って、各発言をミキシングすることができる。スケジューラ22は、発言が実際に行われた時刻に基づいて配信時刻を制御する。

【0027】これにより、図6を参照して説明した実施形態の利点に加え、より厳密に誰が最初に発言を行なったかを決定することができるので、早いもの順で行なうサービス、例えばクイズのようなサービスも公正に行なうことが可能になる。本発明の音声対話システムは、上記端末それぞれが、論理上の人物像（いわゆるアバタ）を有し、上記スケジューラが、論理上の人物像空間内における該人物像の論理上の位置関係、距離、周囲環境、視線方向、視界、姿勢のうちの少なくとも1つに基づいて音声信号の配信を制御するものであってもよい。

【0028】図9は、この態様を示す原理図である。図9に示すように、それぞれの端末A, Bには、それら各端末A, Bのユーザ（すなわち各発言者）を表わす論理上の人物像（いわゆるアバタ）の位置、視線方向、視界等を管理するアバタ管理情報が格納されるアバタ管理情報メモリ17A, 17B、端末A, Bのアバタが共有する論理上の世界を管理する世界管理情報メモリ18A, 18Bが備えられている。各端末A, Bのユーザは、自分自身を表わすアバタの位置、視線方向等を制御するこ

とができる。

【0029】各端末A、Bは、発言を送信するとともに、自分自身のアバタの移動方向、視線の移動方向等を送信し、サーバ20は、発言はサーバ20内のバッファメモリ21A、21Bに記憶し、さらに、そのアバタの論理的な位置、視線方向、視界等を管理するアバタ管理情報メモリの内容を逐次更新する。サーバ20には、各端末A、Bと同様の世界管理情報メモリ30と、アバタどうしがどこまで離れていても話ができるかを示す最大到達距離を格納する最大到達距離メモリ31が備えられている。スケジューラ22において、各端末A、Bに発言を順次送信する際に、ある発言者(＝アバタ)から見えない(視界外にいる)アバタからの発言は送信しない、もしくは、優先順位を落として送信するなどのスケジューリングを行なう。

【0030】これにより、本発明の基本的な利点に加え、さらに、見えてない発言者から話し掛けられるという不都合や、あまりに距離の離れている発言者からの発言は聞こえなくするなどの制御を行なうことが可能となる。さらに、このように、聞く必要のない発言の配信を制限することで、不必要な通信量を削減することも可能となる。

【0031】本発明の音声対話システムは、上記端末それぞれが論理上の人物像(アバタ)を有し前記論理上の人物像空間内における、該人物像の位置関係、距離、周囲環境、視線方向、視界、姿勢のうちの少なくとも1つに基づいて音声信号に変更を加えるエフェクタ手段を備えたものであることが好ましい。図10は、この態様を示す原理図である。

【0032】図10に示すように、それぞれの端末A、Bより発言を送信するとともに、各発言者(各アバタ)の移動方向、視線の移動方向等を送信し、発言はサーバ20内のバッファメモリ21A、21Bに記憶し、さらに、各アバタの論理的な位置、視線方向、視線等を管理するアバタ管理情報メモリの内容を逐次更新する。バッファリングされた発言を、それぞれの端末に配信する際に、論理的に非常に離れた発言者からの発言は配信しないようにし、それ以外の場合は、エフェクタ32により、発言者と聴取者との論理的な位置関係、それぞれの視線方向を考慮し、距離が離れるに従い音量を小さくして加工する。また、方位に従って、左右の音量比、位相、音質(フィルタの特性)を変化させる。さらに、発言者と聴取者の間に障害物がある場合には、間接的に音が伝わっていることを表現するために、反響音を付加したり、音質(フィルタの特性)を変更する。このようにして変更した音声信号は、ステレオ音声信号として各出力バッファエリアL、Rに一旦格納する。

【0033】このように加工した、各端末からの発言を適当にスケジューリングを行い、順次各端末に配信し、あるいは同時刻の各発言をミキシングして、各端末に配

信する。これにより、発言者に距離に対応した音量、発言者のいる方位に対応した方向から音がしているように聞こえるなど、より現実感を増すことができる。

【0034】本発明の音声対話システムにおいて、上記サーバが、前記複数の端末のうち対話を行なう対話グループに属する端末を登録する1以上のグループ管理テーブルを備え、上記スケジューラが、受信した音声信号が該音声信号を発信した端末と同一の対話グループに属する端末に配信されるように、音声信号の配信を制御するものであることも好ましい態様である。

【0035】図11は、この態様を示す原理図である。図11に示すように、各受信グループ毎の受信グループテーブルと各発信グループ毎の発信グループテーブルとからなるグループ管理テーブル33を設けて、各グループごとの受信グループテーブルに、そのグループに所属しているすべてのユーザのユーザIDを記述しておく。さらに、各グループごとの発信グループテーブルに、そのグループに所属しているユーザのグループIDを記述しておく。各端末A、Bから送信されてきた発言は、サーバ20内のバッファメモリ21A、21Bに記憶され、スケジューラ22によって適当に順序付けがなされて、同一グループに所属するすべての端末に順次配信される。あるいは、前述したように、ミキサによって同一グループに所属するそれぞれの端末用に時刻を揃えてミキシングして、それぞれの端末に配信してもよい。

【0036】本発明の音声対話システムにおいて、上記端末が、自分が送信する音声信号を受信する受信先の端末を指定する複数の発信モードのうちのいずれか1つの発信モードを選択する発信モード選択手段を備えたものであることが好ましい。図12は、この態様を示す原理図である。図12に示すように、あらかじめ、発信モード選択手段19A、19Bによって、発言する対象を、同一グループ、または、複数のグループに対する同報、または、特定のユーザ(ないしユーザ群)の中から選択しておく。

【0037】もし、同一グループのみならば、通常の会話モードであり、図11を参照して説明した処理を行なう。もし、複数グループに対する同報が、サーバ20に指示されたなら、グループ管理テーブル33のそのユーザの発信グループテーブル35に同報したい受信グループIDを追加しておく。各端末A、Bから送信されてきた発言は、そのユーザの発信グループテーブル35を参照することで受信グループIDが決定できるので、これに従って、各受信グループの全ての所属ユーザに、発言を配信する。

【0038】もし、特定のユーザ(ないしユーザ群)に対しての発言がサーバ20に指示されたなら、指定されたユーザID(ないしユーザ群のID)を、グループ管理テーブル33の、そのユーザの発信グループテーブルに追加する。各端末A、Bから送信されてきた発言は、

そのユーザの発信グループテーブル35を参照することで受信グループIDとユーザIDが決定できるので、これに従って、各受信グループの全ての所属ユーザと、すべてのユーザ（ないしユーザ群）に発言を配信する。

【0039】本発明の音声対話システムにおいて、上記端末が、自分が受信する音声信号の送信元の端末を指定する複数の聴取モードのうちのいずれか1つの聴取モードを選択する聴取モード選択手段を備えることも好ましい態様である。図13は、この態様の原理図である。図13に示すように、あらかじめ、聴取モード選択手段101Bによって、聴取する対象を同一グループ、または、複数のグループを同時に聴取するか、または、特定のユーザの発言も聴取するのかが選択しておく。

【0040】もし、同一グループのみならば、通常の会話モードであり、図11を参照して説明した処理を行なう。もし、複数グループ同時の聴取がサーバ20に指示されたなら、グループ管理テーブル33の、聴取したいグループの受信グループテーブル34のエントリに自分のユーザIDを追加しておく。各端末A、Bから送信されてきた発言は、そのユーザの発信グループテーブル35を参照することで受信グループIDを決定する。これに従って、受信グループテーブル34に登録されている全ての所属ユーザに発言を配信する。この結果、受信グループテーブル34に追加されたユーザIDにも発言が配信される。

【0041】もし、特定のユーザ（ないしユーザ群）に対しての聴取がサーバ20に指示されたなら、指定されたユーザ（ないしユーザ群）の発信グループテーブル30に自分のユーザIDを追加する。各端末A、Bから送信されてきた発言は、そのユーザの発信グループテーブル35を参照することで受信グループIDおよびユーザIDを決定する。これに従って、受信グループテーブル34に登録されている全ての所属ユーザおよび指定されたユーザに、発言を配信する。この結果、発信グループテーブル35に追加されたユーザIDにも発言が発信される。

【0042】本発明の音声対話システムにおいて、上記端末が、サーバに向けて、音声信号の発信元である端末のうちの、自分あてに送信される音声信号の受信を拒否する端末を指定する割込禁止通知を行なう割込禁止通知手段を備え、上記サーバが、割込禁止通知により指定された送信元の端末から受信した音声信号の送信先が割込禁止通知を行なった端末であった場合に、その送信先の端末へのその音声信号の転送を取り止めるものであることも好ましい態様である。

【0043】図14は、この態様の原理図である。図14に示すように、それぞれの端末において、割込禁止をサーバ20に通知することで、ユーザの指定したグループ、ユーザから以外の発言をサーバより発信しないようにする。サーバ20では、端末より割込禁止が通知され

たなら、受信グループに対する割込禁止ならば、受信グループテーブル34の該当グループの割込禁止フラグをONにし、自端末に対する割込禁止ならば、発信グループテーブル35の該当端末IDの割込禁止フラグをONにする。あるユーザから発言モードの指定があった場合、受信グループテーブル34の割込禁止フラグ、発信グループテーブル35の該当する端末IDの割込禁止フラグを確認し、フラグがONならば、発信グループテーブル35への登録を拒否する。

【0044】本発明の音声対話システムにおいて、上記対話グループに対応してパスワードが付されており、上記端末が、パスワードを入力して前記サーバに送るパスワード入力手段を備え、上記サーバが、端末から送られてきたパスワードを検査して正しいパスワードを送信してきた端末を、該パスワードに対応する対話グループに加入させるものであることも好ましい態様である。

【0045】図15は、この態様を示す原理図である。図15に示すように各端末には、パスワード入力手段102A、102Bが備えられており、サーバ20には、セキュリティ検査手段37、ユーザごとの特権レベルテーブル36、受信グループテーブル34、発信グループテーブル35の各エントリに、特権レベルメモリ、パスワードメモリが付加されている。特権レベルは、各ユーザがどのレベルのグループの対話に参加が許されるかを示す指標であり、パスワードは、そのパスワードを知っているユーザのみ対話に参加できるようにするためのものである。

【0046】ここでは、ユーザからの要求に従って、ユーザの所属しない他のグループの発言を聴取するために他のユーザの発信グループテーブル34や、他のグループの受信グループテーブル35を書き換えようとする際に、そのユーザの特権レベル、および、パスワードをセキュリティ検査手段39によって検査し、それが正当な要求である場合のみに、書き換えが行なわれる。

【0047】本発明の音声対話システムにおいて、上記端末が、話者と聴取者を表示する表示手段を備えることも好ましい態様である。図16は、この態様を示す原理図である。各端末A、Bには、表示手段105A、105B、表示手段105A、105Bに表示されている複数のユーザ（アバタ）の中の話しかけようとするユーザを指定する座標入力手段104A、104B、表示手段105A、105Bに描画するための描画手段103A、103Bが備えられている。表示手段105A、105Bには、誰が誰に話しかけているかが明示される。

【0048】このように、話し掛ける相手を指示する手段と、誰が誰に話し掛けているかを明示する手段を付加することで、話し掛けるときに、話し掛けたい相手を指示しながら発言することで、発言に相手のユーザIDも付加してサーバ20に送信し、サーバ20はそのデータに、話者のユーザIDも付加して、各端末に配信する。

【0049】これにより、各端末A、Bでは、サーバ20から配信された発言に付加されている、話者のユーザID、相手のユーザIDを読み取り、話者のイラスト（アバタ）、および、相手のイラストを変化させる。本発明の音声対話システムにおいて、上記サーバが、音声信号の音質を変化させるエフェクタを備えることも好ましい態様である。

【0050】図17は、この態様を示す原理図である。図17に示すように、ユーザが希望がする場合は、音質を変化させるか否かのモードを設定するモード設定手段106A、106Bにより発言に匿名フラグを付加してサーバ20に送信する。サーバ20では、匿名フラグの付加された発言を受信した場合は、発言をエフェクタ38に通すことで、音響を変化させて、話者を特定できないようにした後、各端末に配信する。

【0051】本発明の音声対話システムにおいて、上記サーバが、所定の効果音を表わす効果音信号を格納する効果音バッファを備え、上記スケジューラが、送信されてきた音声信号の配信を制御するとともに、前記効果音バッファに格納された効果音信号の配信を制御するものであることも好ましい態様である。図18は、この態様を示す原理図である。

【0052】サーバ20には、効果音を格納する効果音バッファ39、その効果音を管理する効果音管理情報メモリ、その効果音の反復間隔を格納する反復間隔メモリ41が備えられている。ここでは、ユーザの発言以外の、例えば、商店の店内放送や、駅のアナウンスなどの効果音のPCMデータと、その論理的位置、反復間隔をサーバ20に登録すると、サーバ20では、これらの効果音を、指定された間隔において反復して配信するように、スケジューリング、または、ミキシングを行なう。これにより、効果音も、発言と同様に、発言の合間に、各端末に配信される。

【0053】本発明の音声対話システムにおいて、上記サーバもしくは端末が、音声信号をテキストデータに変換する音声認識手段を備えることも好ましい態様である。図19は、この態様を示す原理図である。図19に示すように、端末から送られてきた発言を各端末に配信する際に、音声出力手段を持たない端末に対して発言を配信する場合は、サーバ20に付加した音声認識手段42を用いることで、バッファリングされている音声のPCMデータを、テキストに変換する。これにより、音声出力手段を持たない端末においては、テキストにて発言を表示することができる。端末側に音声認識手段107Bをもってもよい。

【0054】図20は、もう1つの態様を示す原理図である。図20に示すように、音声入力手段を持たない端末から、テキストにより受信した発言を、サーバ20に付加した音声合成手段43を用いてPCMデータに変換し、音声出力手段を持つ端末に対しては、音声によって

データを配信する。音声出力手段を持たない端末に対しては、テキストにより端末より受信した発言はそのまま配信し、音声により端末より受信した発言は、音声認識手段42によってテキストに変換されて配信される。

【0055】図21は、さらにもう1つの態様を示す原理図である。図21に示すように、ユーザの使用言語をあらかじめサーバ20に通知し使用言語テーブル45に格納しておき、端末より受信した発言の使用言語と、配信する端末の使用言語が不一致の場合、発言を音声認識手段42によりテキストに変換したのち、翻訳手段44により、言語の翻訳を行い、さらに、翻訳結果のテキストを音声合成手段43によって音声に変換する。これにより、使用言語の違いを気にすることなく会話することができる。

【0056】図22は、もう1つの態様を示す原理図である。図22に示すように、受信した発言を音声認識手段43によりテキストに変換し、変換されたテキストと語句テーブル47にあらかじめ登録されている指定語句とを、サーバ20に付加された検索手段46により検索し、もし、指定語句が存在したら、これを削除する。

【0057】これにより不適切な表現が聴取者に送信されるのを避けることができる。本発明の音声対話システムにおいて、上記サーバと端末がアナログ回線で接続され、アナログ信号としての音声信号の始端および終端に音声信号の始端、終端を識別するためのトーン音を発生させるとともに、受信したトーン音を検出する、それぞれ、トーン音発生手段およびトーン音検出手段を備えたことも好ましい態様である。この態様では、図23に示す様に、端末、サーバに図示の各手段が備えられており、この図23に示すように、回線1から送られて来たデータがアナログの音声データならば、そのままスピーカに出力し、デジタルデータならばスピーカへの出力を禁止する、データスイッチ手段を端末に付加する。サーバ20よりアナログデータ開始コードが送られてきたならば、データスイッチを切り替えてスピーカ出力をONにし、データ通信モードをアナログにする。この状態で、あらかじめ定められた時間以上の無音期間に続いて識別音が入力されたら、データスイッチを切り替えてスピーカスイッチをoffにし、データ通信モードをデジタルに設定する。

【0058】これにより、音声情報はアナログにより伝送するので、電話回線のような低速な回線でも利用できる。本発明の音声対話システムにおいて、上記サーバが、前記端末どうしの対話を記憶する記憶手段を備えることも好ましい態様である。その場合に、上記サーバが、端末どうしの対話を記憶手段に記憶するにあたりその対話に付随する付随情報を付加して記憶するものであることがさらに好ましい。

【0059】図24は、この態様を示す原理図である。図24に示すように、各端末からの発言を、各端末に配

信するとともに、大容量記憶装置50に、発言時刻、発言者、発言相手の情報を付加して記憶する。そして、端末からの要求により、会話中、および、会話終了後に、その会話を再生できる。

【0060】図24に示す態様において、上記サーバが、上記付属情報に基づいて、記憶手段に記憶された対話中の、所望の対話部分を検索する検索手段を備えることがさらに好ましい。図25は、この態様を示す原理図である。図25に示すように、サーバに検索手段51を付加することにより、発言時刻、発言者、発言相手をキーワードとして、会話の記録された大容量記憶装置より検索することにより、該当する発言のみを再生することができる。

【0061】さらに、対話の話速を変換する話速変換手段を備えることが好ましい。図26は、この態様を示す原理図である。図26に示すように、大容量記憶装置50に記録された会話を再生する際に、実際に発言が行なわれた時間タイミングで再生しても良いが、好ましくは、話速変換手段52によって、空白部分を省略して発言を時間順に連続して再生しても良いし、発言を実際の時間より短縮して再生しても良い。

【0062】

【発明の実施の形態】以下、本発明の実施形態について説明する。図27は、本発明の第1実施形態を示すブロック図である。本実施形態においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクロホン（以下、マイクと略称する）からの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつパケットにし、ユーザIDと端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したパケットをサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0063】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリ（メモリA、メモリB）に保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。

【0064】スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信

時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0065】図28は、本発明の第2実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、サーバに同時発言者数メモリが付加される。本実施形態においては、マウスを用いて、発言要求と発言終了を通知する。また、発言許可通知手段はディスプレイにより実現される。つまり、マウスにより発言要求ボタンをクリックすることで、サーバに対して発言要求を行う。サーバから許可が与えられたら、発言要求ボタンの表示を、「発言許可」に切り替える。これ以降は、第1実施形態と同様に、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行ない、このデジタルデータを、例えば256バイトづつパケットにし、ユーザIDと端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したパケットをサーバに送信するとともに、発言終了通知をサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0066】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0067】同時発言者数管理手段はCPUと同時発言者数メモリによって実現され、端末より、発言要求がきたならば、この同時発言者数メモリに保持されている数値を確認し、1以上なら、その端末に許可通知を与えるとともに、数値を-1とする。もし、数値が既に0となっていたならば、どれかの端末から発言終了通知が来るまで待つ。

【0068】端末側では、許可通知を受信したなら、ディスプレイに発言が許可されたことを表示し、ユーザに発言を行っても良いことを知らせる。ユーザはこの表示を確認の後、マウスボタンを押しながら発言を開始し、発言完了後にマウスボタンを放すことで、発言終了通知をサーバに対して行う。サーバ側では、発言終了通知を

受信したならば、同時発言者数メモリに保持されている数値を+1する。

【0069】図29は、本発明の第3実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、無音部検出手段、最大無音量メモリ、最短無音時間メモリが付加される。無音部検出手段は、CPUによって実現され、バッファメモリ内のPCMデータの絶対値を調べることにより、音量の大きさを計測し、音量が予め設定された値“最大無音量”より小さい期間が、予め設定された時間“最短無音時間”より長ければ、最初に最大無音量を下回った部分で、発言を分割する。

【0070】サーバ制御部は、スケジューラにより指示されたバッファメモリ上の発言を、各端末に配信するが、このとき、無音部検出手段を呼び出し、無音部と判定されたら、発言の配信をそこでいったん打ち切り、次に最大無音量を越える部分（無音区間をスキップする）を発言の先頭とし、その対応する時刻をバッファメモリに記録する。

【0071】その他の処理は、第1実施形態と同様である。端末側のもう一つの実施形態を、図29のPC-Bの端末に示す。この実施形態では、マウスボタンにより、発言の開始と終了を指示するのではなく、上記無音部の検出法と同様に、マイクからの入力に対して常時A/D変換を行ない、そのPCM出力データによって、発言の開始点、終了点を検出する。

【0072】CPUによって無音検出手段を実現し、A/D変換結果のPCM出力データの絶対値を調べることにより、音量の大きさを計測し、音量が予め設定された値“最大無音量”より大きくなったなら、そこから発言がはじまったものとして、サーバに順次データを送信する。そして、PCM出力データの音量が、“最大無音量”より小さい期間が、予め設定された時間“最短無音時間”より長ければ、最初に最大無音量を下回った部分が発言の終了点とみなし、サーバへデータ終了フラグを付加したパケットを送信する。そして、引き続き、A/D変換結果のPCM出力結果を監視して、次の発言の開始点を待つ。

【0073】図30は、本発明の第4実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、各バッファメモリに受信開始時刻メモリが付加される。サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このとき、各端末からの発言の先頭パケットを受信した時刻を受信開始時刻メモリに記録しておく。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。

【0074】スケジューラはCPUにより実現され、端

末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信開始時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。もちろん、図29に示す実施形態と同様に、サーバ側に、無音部検出手段を付加しても良いし、端末側もマウスを使わず、無音部検出手段によって、発言の開始点、終了点を検出しても良い。

【0075】図31は、本発明の第5実施形態を示すブロック図である。本実施形態は、図30に示す第4実施形態と同様に構成されるが、端末が接続されたとき最初に実行される内部時間補正值計算部と、内部時間補正值メモリが付加されることと、受信開始時刻メモリに代わって、補正済受信開始時刻メモリを持つことが異なる。

【0076】内部時間補正值計算部は、CPUにより実現される。端末が新たに接続されると、まず、端末に対して、端末のローカルな時刻を内部時計より読み取り、その時刻 T_1 をサーバに対して送信するよう指示する。この指示を出した時刻を T_s とする。サーバでは、この指示に回答して、端末のローカルな時刻が端末より報告されるのを待つ。端末から報告が得られたら、この時刻を T_e とする。報告された端末の内部時計とサーバの内部時計との差分は、内部時間補正值 $T_d = T_s + (T_e - T_s) / 2 - T_1$ となる。この値をすべての端末について計算して、各端末用バッファの内部時間補正值メモリに記録する。

【0077】各端末の制御部はCPUにより実現され、発言が開始された内部時刻 T_x を内部時計により計測し、その時刻を発言データパケットに付加してサーバに送信する。サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このとき、各端末からの発言に付加されている内部時刻 T_x を読み取り、補正済受信開始時刻=内部時刻 T_x -内部時間補正值 T_d を計算し、この時刻を補正済受信開始時刻メモリに記録する。

【0078】スケジューラはCPUにより実現され、端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、補正済受信開始時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。これにより、スケジューラは、発言が実際に行なわれた順に従って、各発言を順次、すべての端末に送信することができる。

【0079】もちろん、図29に示す実施形態と同様に、サーバ側に、無音部検出手段を付加しても良いし、端末側もマウスを使わず、無音部検出手段によって、発言の開始点、終了点を検出しても良い。図32は、本発

明の第6実施形態を示すブロック図である。本実施形態においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつパケットにし、ユーザIDと端末IDを付加して、サーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0080】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに一時保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、出力バッファメモリ上のミキシングされた発言を各端末に配信する。ミキサーはサーバのCPUにより実現され、各バッファメモリに保持されている発言データの時刻の同じ部分を加算することでミキシングし、ゲイン調整をしたのち出力バッファメモリに保持する。もちろん、DSPなどにより実現しても良い。

【0081】同時刻の決定方法としては、例えば、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録しておき、処理しようとする時刻のデータを、バッファの先頭位置からのオフセットより求めることで、各バッファ内の同時刻のデータを取り出すことができる。ゲイン調整の方法としては、例えば、加算結果がPCMの最大値を越えたならば、PCMの最大値にクリッピングする。または、加算結果がPCMの最大値を越えたなら、PCMの最大値になるよう正規化し、その時の係数をゲイン係数として保持し、以降の加算結果に対しても適用する。ゲイン係数は、時定数により1に減衰するように制御する。

【0082】また、端末は、図29に示す実施形態で示したように、無音部検出手段を設けることで、マウスを用いずに、マイク入力のシグナルレベルにより、自動的に発言の開始、終了を検出するよう構成しても良い。図33は、本発明の第7実施形態を示すブロック図である。本実施形態は、図32に示す第6実施形態と同様に構成されるが、各バッファメモリに受信開始時刻メモリが付加される。この例においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつパケットにし、ユーザIDと端末IDを付加して、サーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0083】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに一時保

持する。このとき、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、出力バッファメモリ上のミキシングされた発言を各端末に配信する。

【0084】ミキサーはCPUにより実現され、各バッファメモリに保持されている発言データのバッファの先頭からのオフセットと発言の受信時刻より、処理したい時刻のデータを取り出すことができる。各バッファの同時刻のデータを加算することでミキシングし、ゲイン調整をしたのち出力バッファメモリに保持する。ゲイン調整の方法としては、例えば、加算結果がPCMの最大値を越えたならば、PCMの最大値にクリッピングする。または、加算結果がPCMの最大値を越えたなら、PCMの最大値になるよう正規化し、その時の係数をゲイン係数として保持し、以降の加算結果に対しても適用する。ゲイン係数は、時定数により1に減衰するように制御する。

【0085】これらの、加算処理、ゲイン調整処理は、DSPなどにより実現しても良い。また、端末は、図29に示す第3実施形態で示したように、無音部検出手段を設けることで、マウスを用いずに、マイク入力のシグナルレベルにより、自動的に発言の開始、終了を検出するよう構成しても良い。図34は、本発明の第8実施形態を示すブロック図である。

【0086】本実施形態は、第7実施形態と同様に構成されるが、端末が接続されたとき最初に実行される内部時間補正值計算部と、内部時間補正值メモリが付加されることと、受信開始時刻メモリに代わって、補正済受信開始時刻メモリを持つことが異なる。内部時間補正值計算部は、CPUにより実現される。端末が新たに接続されると、まず、端末に対して、端末のローカルな時刻を内部時計より読み取り、その時刻 T_1 をサーバに対して送信するよう指示する。この指示を出した時刻を T_s とする。サーバでは、この指示に応答して、端末のローカルな時刻が端末より報告されるのを待つ。端末から報告が得られたら、この時刻を T_e とする。報告された端末の内部時計とサーバの内部時計との差分は、内部時間補正值 $T_d = T_s + (T_e - T_s) / 2 - T_1$ となる。この値をすべての端末について計算して、各端末用バッファの内部時間補正值メモリに記録する。

【0087】各端末の制御部はCPUにより実現され、発言が開始された内部時刻 T_x を内部時計により計測し、その時刻を発言データパケットに付加してサーバに送信する。サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このとき、各端末からの発言に付加されている内部時刻 T_x を読み取り、補正済受信開始時刻=内部時刻 T_x -内部時間補正值 T_d を計算し、この時刻を補正済受信開

始時刻メモリに記録する。

【0088】ミキサーはCPUにより実現され、各バッファメモリに保持されている発言データのバッファの先頭からのオフセットと発言の補正済受信時刻より、処理したい時刻のデータを取り出すことができる。各バッファの同時刻のデータを加算することでミキシングし、ゲイン調整をしたのち出力バッファメモリに保持する。ゲイン調整の方法としては、例えば、加算結果がPCMの最大値を越えたならば、PCMの最大値にクリッピングする。または、加算結果がPCMの最大値を越えたなら、PCMの最大値になるよう正規化し、その時の係数をゲイン係数として保持し、以降の加算結果に対しても適用する。ゲイン係数は、時定数により1に減衰するように制御する。

【0089】これらの、加算処理、ゲイン調整処理は、DSPなどにより実現しても良い。また、端末は、図29の実施形態で示したように、無音部検出手段を設けることで、マウスを用いずに、マイク入力のシグナルレベルにより、自動的に発言の開始、終了を検出するよう構成しても良い。図35は、本発明の第9実施形態を示すブロック図である。

【0090】本実施形態は、図30を参照して説明した実施形態と同様に構成されるが、サーバと端末にアバタ管理情報メモリと世界管理情報メモリが、サーバに最大到達距離メモリが付加されることが異なる。本実施形態では、各端末のユーザは、仮想的に3次元座標と、視線方向、視界を持ち、これらの情報が、アバタ管理情報として、各端末と、サーバの両方で保持される。ユーザがマウス、キーボードなどにより、移動や視線の移動を指示すると、端末のアバタ管理情報の該当する情報が更新される。また、これらのアバタ管理情報が変更されると、端末より変更通知がサーバに送信され、サーバ側の該当する情報も同様に更新される。世界管理情報は、仮想的な3次元世界の中に、どのような建物がどこにあるかを示す情報であり、ポリゴンデータである。

【0091】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このとき、各端末からの発言の先頭バケットを受信した時刻を受信開始時刻メモリに記録しておく。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。さらに、制御部は、端末よりアバタ情報更新通知があったなら、該当するアバタ情報を通知に従って更新する。

【0092】スケジューラはCPUにより実現され、各端末ごとに配信すべき発言とその順序を検討する。他の端末から送信されてきて、まだ配信されていない発言があれば、その端末のアバタ管理情報とある発言者のアバタ管理情報と、世界管理情報から、位置関係と視界を計

算し、2者間の距離が、あらかじめ設定された最大到達距離以上ならば、その発言は聞こえないものとして送信しない。また、視界内に入っていないユーザからの発言も、その発言は聞こえないものとして送信しない。それ以外の発言については、図30を参照して説明した実施形態と同様に、古いものから順に配信するよう制御部に指示する。

【0093】もちろん、図29を参照して説明した実施形態と同様に、サーバ側に、無音部検出手段を付加しても良いし、端末側もマウスを使わず、無音部検出手段によって、発言の開始点、終了点を検出しても良い。図36は、本発明の第10実施形態を示すブロック図である。本実施形態は、図32を参照して説明した実施形態と同様に構成されるが、出力バッファエリアがLRの2チャンネルとなり、サーバと端末にアバタ管理情報メモリと世界管理情報メモリが、サーバにエフェクタと最大到達距離メモリが付加されることが異なる。

【0094】本実施形態では、各端末のユーザは、仮想的に3次元座標と、視線方向、視界を持ち、これらの情報が、アバタ管理情報として、各端末と、サーバの両方で保持される。ユーザがマウス、キーボードなどにより、移動や視線の移動を指示すると、端末のアバタ管理情報の該当する情報が更新される。また、これらのアバタ管理情報が変更されると、端末より変更通知がサーバに送信され、サーバ側の該当する情報も同様に更新される。世界管理情報は、仮想的な3次元世界の中に、どのような建物がどこにあるかを示す情報であり、ポリゴンデータである。

【0095】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このとき、各端末からの発言の先頭バケットを受信した時刻を受信開始時刻メモリに記録しておく。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、端末よりアバタ情報更新通知があったなら、該当するアバタ情報を通知に従って更新する。さらに、制御部は他の端末から送信されてきて、まだ配信されていない発言があれば、その端末のアバタ管理情報とある発言者のアバタ管理情報と、世界管理情報から、位置関係と視界を計算し、2者間の距離が、あらかじめ設定された最大到達距離以上ならば、その発言は聞こえないものとして配信しない。また、視界内に入っていないユーザからの発言も、その発言は聞こえないものとして送信しない。

【0096】エフェクタはCPUにより実現され、上記以外の発言について、発言者と聴取者との論理的な位置関係、それぞれの視線方向を考慮し、距離が離れるに従い音量を小さくしく加工する。また、方位に従って、左右の音量比、位相、音質（フィルタの特性）を変化させる。さらに、発言者と聴取者の間に障害物がある場合に

は、間接的に音が伝わっていることを表現するために、反響音を付加したり、音質（フィルタの特性）を変更する。

【0097】このようにして加工した、各端末からの発言を、適当にスケジューリングを行って、順次各端末に配信し、あるいは同時刻の各発言をミキシングして、各端末に配信する。もちろん、図29を参照して説明した実施形態と同様に、サーバ側に、無音部検出手段を付加しても良いし、端末側もマウスを使わず、無音部検出手段によって、発言の開始点、終了点を検出しても良い。

【0098】図37は、本発明の第11実施形態を示すブロック図、図38は、グループ管理テーブルの一例を示す図である。本実施形態は、例えば、図27に示す実施形態と同様に構成されるが、サーバにグループ管理テーブルが付加されることが異なる。この例においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつバケットにし、ユーザIDと端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したバケットをサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0099】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を、グループ管理テーブルに登録されている端末IDが示す各端末に配信する。

【0100】スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭バケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0101】ここでは、図27と同様のスケジューラを用いた実施形態を示したが、図32と同様のミキサを用いた実施形態も同様に構成できる。図39は、本発明の第12実施形態を示すブロック図、図40は、発信グループテーブル及び受信グループテーブルの一例を示す図である。本実施形態は、例えば、図37に示す実施形態と同様に構成されるが、端末に発言モード選択手段が、サーバに発信グループテーブルと受信グループテーブルが付加されることが異なる。

【0102】この例においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつバケットにし、ユーザIDと発信端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したバケットをサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0103】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの発信端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を、その発信端末IDを参照して、発信グループに登録されている端末IDおよび受信グループIDが示す受信グループに登録されている端末IDが示す各端末に配信する。

【0104】スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭バケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0105】発言モード選択手段は、CPUとマウスによって構成され、ユーザの希望するモードを示すアイコンをクリックすることで、モードを指示することができる。ここでは、図27と同様のスケジューラを用いた実施形態を示したが、図32と同様のミキサを用いた実施形態も同様に構成できる。図41は、本発明の第13実施形態を示すブロック図、図42は、発信グループテーブル及び受信グループテーブルの一例を示す図である。

【0106】本実施形態は、例えば、図37に示す実施形態と同様に構成されるが、端末に聴取モード選択手段が、サーバに発信グループテーブルと受信グループテーブルが付加されることが異なる。この例においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつバケットにし、ユーザIDと発信端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したバケットをサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通し

て、これを音声に変換し、スピーカより鳴らす。

【0107】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの発信端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を、その発信端末IDを参照して、発信グループに登録されている端末IDおよび受信グループIDが示す受信グループに登録されている端末IDが示す各端末に配信する。

【0108】スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0109】聴取モード選択手段は、CPUとマウスによって構成され、ユーザの希望するモードを示すアイコンをクリックすることで、モードを指示することができる。ここでは、図27と同様のスケジューラを用いた実施形態を示したが、図32と同様のミキサーを用いた実施形態も同様に構成できる。図43は、本発明の第14実施形態を示すブロック図、図44は、発信グループテーブル及び受信グループテーブルの一例を示す図である。

【0110】本実施形態は、図41に示す実施形態と同様に構成されるが、サーバの受信グループテーブルと発信グループテーブルに割込禁止フラグが付加される。サーバでは、端末より割込み禁止が通知されたなら、受信グループに対する割込禁止ならば、受信グループテーブルの該当グループの割込禁止フラグをONにし、自端末に対する割込禁止ならば、受信グループテーブルの自端末IDの割込禁止フラグをONにする。あるユーザから、発言モードの指定があった場合、受信グループテーブルの割込禁止フラグ、受信グループテーブルの該当する端末IDの割込禁止フラグを確認し、フラグがONならば、発信グループテーブルへの登録を拒否する。

【0111】また、これと、後述する、図47に示す実施形態を組合せて、発言モードの指定を行う場合に、特権レベル値が、受信グループの値、または、端末IDの値より高ければ、割込禁止フラグの如何に拘らず、その受信グループ、または、その端末に対して、発言を行えるようにしても良い。このように構成された、本発明の第15実施形態を、図45、図46に示す。詳細説明は省略する。

【0112】図47は、本発明の第16実施形態を示す

ブロック図、図48は、発信グループテーブル及び受信グループテーブルの一例を示す図である。本実施形態は、例えば、図41に示す実施形態と同様に構成されるが、端末にパスワード入力手段が、サーバにセキュリティ検査手段、ユーザごとの特権レベルテーブル、発信グループテーブル、発信グループテーブルの各エントリに特権レベルメモリ、パスワードメモリが付加されることが異なる。

【0113】この例においては、マウスを用いて、発言の開始と終了を指示する。つまり、マウスボタンを押している期間中、マイクからの入力をサンプリングして、A/D変換を行い、このデジタルデータを、例えば256バイトづつパケットにし、ユーザIDと発信端末IDを付加して、サーバに送信する。マウスボタンを放したなら、データ終了フラグを付加したパケットをサーバに送信する。サーバより配信される他の端末からの発言は、受信しだい、D/A変換器に通して、これを音声に変換し、スピーカより鳴らす。

【0114】サーバ制御部はCPUによって実現され、各端末から送信されてきたデータの発信端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。バッファメモリは、接続されている端末数と同数用意される。また、制御部は、スケジューラにより指示されたバッファメモリ上の発言を、その発信端末IDを参照して、発信グループに登録されている端末IDおよび受信グループIDが示す受信グループに登録されている端末IDが示す各端末に配信する。

【0115】スケジューラはCPUにより実現され、適当な方針により、バッファメモリに保持されている発言データを、各端末に配信する順序を決定する。例えば、各端末からの発言の先頭パケットを受信した時刻をバッファメモリに記録しておく。端末から送信されてきて、まだ配信されていない発言があれば、それを各端末に配信するよう制御部に指示する。ここで、もし、まだ配信されていない発言が複数ある場合は、それらの内、受信時刻のもっとも古いものを選択し、それを送信するよう制御部に指示する。

【0116】聴取モード選択手段は、CPUとマウスによって構成され、ユーザの希望するモードを示すアイコンをクリックすることで、モードを指示することができる。セキュリティ検査手段は、CPUによって実現され、ユーザからの受信テーブル及び発信テーブルの書き換え要求を受けたときに、そのユーザの特権レベル値が、各テーブルのエントリの特権レベル値より小さければ、書き換えを無条件に許す。しかし、特権レベル値が大きい場合は、端末よりパスワードを入力させ、これが、テーブルのエントリと一致した場合のみ、書き換えを許す。

【0117】ここでは、図27と同様のスケジューラを用いた実施形態を示したが、図32と同様のミキサーを

用いた実施形態も同様に構成できる。図49は、本発明の第17実施形態を示すブロック図、図50は、ディスプレイ上に表示されたイラストの例を示す図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、端末に制御部と表示手段と座標入力手段と描画手段が付加される。

【0118】本実施形態では、対話を行なっている相手のユーザは、ディスプレイ上に人間のイラストとして表示され、このイラストをマウス等でクリックしながら話すことで、話す相手を指定する。端末の制御部はCPUによって実現され、マウスの動きを監視して、マウスによってディスプレイ上のどの人間のイラストがクリックされたかを検出し、話す相手特定する。

【0119】描画手段もCPUによって実現され、発言しているユーザに対応するディスプレイ上のイラストに、吹き出しのイラストを付加することで、誰が発言しているかを明示することができる。また、話し掛けられている人のイラストには、呼出音のイラストを付加することで、誰に話しているかを明示できる。自分に対して話し掛けられた場合は、ピッという音とともに、画面左上に呼出音のイラストが表示される。

【0120】その他の処理は、図27に示す第1実施形態と同様である。図51は、本発明の第18実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、サーバにエフェクタが、端末に制御部とモード設定手段が付加される。モード設定手段および制御部はCPUにより実現され、ユーザがアイコンをクリックすることで発言を匿名モードに設定した場合、制御部は発言に匿名フラグを付加して、サーバに送信する。

【0121】サーバの制御部は、CPUにより実現され、端末からの発言に匿名フラグが設定されていたならば、エフェクタを用いて、音声波形にフィルタリングを行い、音質を変化させる。エフェクタもCPUにより実現され、FFT、IFFTなどにより音声のスペクトラムを変化させる。

【0122】その他の処理は、第1実施形態と同様である。図52は、本発明の第19実施形態を示すブロック図である。本実施形態は、図30に示す第4実施形態と同様に構成されるが、サーバに効果音メモリと効果音管理情報メモリと反復間隔メモリと受信開始時刻メモリが付加される。

【0123】スケジューラはCPUにより実現され、各端末ごとに配信すべき発言とその順序を検討する。他の端末から送信されてきて、まだ配信されていない発言があれば、古いものから順に配信するよう制御部に指示する。このとき、効果音メモリの内容も、反復間隔ごとに、各端末に配信するように、スケジューリングを行なう。つまり、反復間隔ごとに、端末より届いた発言と同様に受信開始時刻を設定し、スケジューラにより他の発

言と同様に配信する。

【0124】その他の処理は、図30の実施形態と同様である。図53は、本発明の第20実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、サーバに音声認識手段が、端末に文字表示手段が付加される。サーバ制御部はCPUによって実現され、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。このとき、配信しようとする端末が音声出力手段を持たないならば、音声認識手段を用いて、これをテキストに変換した後、配信する。

【0125】音声認識手段はCPUによって実現され、例えば、Windows Sound Systemに付属する音声認識プログラムである。文字表示手段はフレームバッファとディスプレイによって実現され、受信したテキストをディスプレイに表示する。その他の処理は、図27の実施形態と同様である。また、音声認識手段を、端末に付加して、端末側でテキストに変換しても良い。

【0126】図54は、本発明の第21実施形態を示すブロック図である。本実施形態は、図53に示す第20実施形態と同様に構成されるが、サーバに音声合成手段が付加される。サーバ制御部はCPUによって実現され、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。このとき、配信しようとする発言がテキストで、かつ、配信しようとする端末は音声出力手段を持つならば、音声合成手段を用いて、これをPCM音声に変換した後、配信する。

【0127】音声合成手段はCPUによって実現され、例えば、Windows Sound Systemに付属する音声合成プログラムである。その他の処理は、図53に示す実施形態と同様である。図55は、本発明の第22実施形態を示すブロック図である。本実施形態は、図54に示す実施形態と同様に構成されるが、サーバに翻訳手段、使用言語テーブルが付加される。

【0128】端末は、発言をサーバに送信するさいに、その使用言語コードを付加して送信する。サーバ制御部はCPUにより実現され、スケジューラにより指示されたバッファメモリ上の発言を各端末に配信する。このとき、発言の使用言語コードと、使用言語コードに記述されている各端末の使用言語コードが一致しなければ、発言を音声認識手段によりテキストに変換したのち、翻訳手段により翻訳し、さらに、音声合成手段によりPCM音声に変換してのち、端末に配信する。

【0129】翻訳手段はCPUにより実現され、例えば、ATLAS翻訳システムである。その他の処理は、図54に示す実施形態と同様である。図56は、本発明の第23実施形態を示すブロック図である。本実施形態は、図54に示す実施形態と同様に構成されるが、サーバに検索手段と、語句テーブルが付加される。

【0130】サーバ制御部はCPUにより実現され、各

端末から送信されてきたデータの端末IDを参照して、発言をそれぞれの端末用のバッファメモリに保持する。このPCM音声データを音声認識手段を用いてテキストデータに変換する。さらに、検索手段を用いて、このテキスト中に語句テーブルに登録されている語句があるかどうかを検索する。もし、あれば、これを削除または編集する。そして、このテキストデータを音声合成手段を用いてPCM音声データに変換する。そして、スケジューラの指示に従って、この発言を各端末に配信する。

【0131】その他の処理は、図54の実施形態と同様である。図57は、本発明の第24実施形態を示すブロック図である。本実施形態においては、端末からの発言はアナログ音声のまま、サーバに送信され、サーバから各端末に発言を配信する際も、アナログ音声のまま配信される。

【0132】端末とサーバをアナログの電話回線によって接続し、モデムを用いてデジタルデータを送受信するような構成において、発言のようなアナログデータは、デジタル化しないで、アナログのまま通す。アナログデータの終了を通知、検知するために、トーン音検出手段、トーン音発生手段、無音区間検出手段と電話回線とスピーカを接続するかどうかを切り替えるための、スピーカスイッチ手段とマイクと電話回線を接続するかどうかを切り替えるマイクスイッチ手段を端末に備える。

【0133】端末制御部において、マウスが押されたことを検知すると、これ以降のデータがアナログとなることを通知するアナログデータ開始コードを、サーバに送信する。そして、マイク入力を電話回線に流すようにマイクスイッチをONにする。そして、マウスが放されて、発言が終了したなら、マイクスイッチをOFFにし、規定時間の無音期間をおいた後、予め定められたトーン音を、トーン音発生手段によって発生させる。これ以降は、モデムを通してのデジタルデータのやりとりに復帰する。また、サーバからアナログデータ開始コードが送信されてきたら、スピーカスイッチをONにして、アナログ音声をスピーカに出力する。このとき、無音区間検出手段によりアナログ音声を監視し、無音区間が発見されてかつ、その直後の音が、予め定められたトーン音であることがトーン音検出手段によって確認されたなら、スピーカスイッチをOFFにして、それ以降のデータをデジタルデータとして、モデムを通して受信する。

【0134】サーバ制御部において、端末よりアナログデータ開始コードが送信されてきたなら、それ以降のデータをA/D変換してPCMデータに変換した後、それぞれの端末用のバッファエリアに保持する。このとき、無音区間検出手段によりアナログ音声を監視し、無音区間が発見されて、かつ、その直後の音が、予め定められたトーン音であることが確認されたなら、A/D変換によるデータ入力を終了する。また、スケジューラの指示に従って、発言を端末に配信する場合は、まず、アナロ

グデータ開始コードを送信したのち、バッファエリアの内容をD/A変換しながら、電話回線にアナログとして出力する。出力終了後、規定の無音区間をおいたあと、予め定められたトーン音をトーン音発生手段により発生させる。これ以降は、モデムを通してのデジタルデータのやりとりに復帰する。

【0135】信号変換手段は、デジタル情報を変調されたアナログ信号に変換し、またその逆に、変調されたアナログ信号をデジタル信号に変換するものである。スケジューラ、その他の処理は、図27に示す第1実施形態と同様である。図58は、本発明の第25実施形態を示すブロック図である。本実施形態は、図27に示す第1実施形態と同様に構成されるが、サーバにハードディスクが付加される。

【0136】サーバ制御部は、各端末から送信されてきた発言を、バッファエリアに保持するとともに、ハードディスクに、発言データ、発言時刻、発言者、発言相手の情報を記録する。また、端末からの要求により、ハードディスクより、記録されている発言を読みだし、端末に配信する。その他の処理は、図27の実施形態と同様である。

【0137】また、ここでは、発言をPCMデータのまま記録しているが、音声認識手段、音声合成手段を用いることで、テキストに変換してから記録し、再生時に音声に戻すこともできるし、テキストのまま見ることも容易に実現できる。図59は、本発明の第26実施形態を示すブロック図である。本実施形態は、図58に示す実施形態と同様に構成されるが、サーバに検索手段が付加される。

【0138】サーバ制御部は、端末からの要求に応じて、ハードディスクに記録されている発言を検索手段によって検索することで、特定の発言時刻、発言者、発言相手をキーワードにして、該当する発言のみを端末に配信することができる。検索手段は、CPUによって実現される。その他の処理は、図58の実施形態と同様である。

【0139】また、音声認識手段をサーバに付加することで、ハードディスクに記録されている発言を、テキストに変換し、これを検索手段で検索することで、発言内容を特定のキーワードで検索することができ、特定の話題についての発言のみを、端末に配信することもできる。図60は、本発明の第27実施形態を示すブロック図である。

【0140】本実施形態は、図58に示す実施形態と同様に構成されるが、サーバに話速変換手段が付加される。サーバ制御部は、端末からの要求に応じて会話を再現する際に、実際に発言が行なわれた時間タイミングで発言を配信するか、もしくは、時間タイミングは無視して、発生順に空白時間は無視して、次々と発言を配信していくか、もしくは、話速変換手段により、1つの発言

内の空白時間（息継ぎの時間）などを削除してのち端末に配信することにより、発言を実際の時間より短縮して聞けるようにする。

【0141】話速交換手段は、無音区間検出器とCPUにより実現され、発言中の無音部を検出し、これを削除することで、発言時間を短縮することができる。その他の処理は、図58に示す実施形態と同様である。

【0142】

【発明の効果】以上説明したように、本発明の音声対話システムによれば、通信量を適正なレベルに制御することができる。

【図面の簡単な説明】

【図1】本発明の一態様を示す原理図である。

【図2】本発明の一態様を示す原理図である。

【図3】本発明の一態様を示す原理図である。

【図4】本発明の一態様を示す原理図である。

【図5】本発明の一態様を示す原理図である。

【図6】本発明の一態様を示す原理図である。

【図7】本発明の一態様を示す原理図である。

【図8】本発明の一態様を示す原理図である。

【図9】本発明の一態様を示す原理図である。

【図10】本発明の一態様を示す原理図である。

【図11】本発明の一態様を示す原理図である。

【図12】本発明の一態様を示す原理図である。

【図13】本発明の一態様を示す原理図である。

【図14】本発明の一態様を示す原理図である。

【図15】本発明の一態様を示す原理図である。

【図16】本発明の一態様を示す原理図である。

【図17】本発明の一態様を示す原理図である。

【図18】本発明の一態様を示す原理図である。

【図19】本発明の一態様を示す原理図である。

【図20】本発明の一態様を示す原理図である。

【図21】本発明の一態様を示す原理図である。

【図22】本発明の一態様を示す原理図である。

【図23】本発明の一態様を示す原理図である。

【図24】本発明の一態様を示す原理図である。

【図25】本発明の一態様を示す原理図である。

【図26】本発明の一態様を示す原理図である。

【図27】本発明の第1実施形態を示すブロック図である。

【図28】本発明の第2実施形態を示すブロック図である。

【図29】本発明の第3実施形態を示すブロック図である。

【図30】本発明の第4実施形態を示すブロック図である。

【図31】本発明の第5実施形態を示すブロック図である。

【図32】本発明の第6実施形態を示すブロック図である。

【図33】本発明の第7実施形態を示すブロック図である。

【図34】本発明の第8実施形態を示すブロック図である。

【図35】本発明の第9実施形態を示すブロック図である。

【図36】本発明の第10実施形態を示すブロック図である。

【図37】本発明の第11実施形態を示すブロック図である。

【図38】グループ管理テーブルの一例を示す図である。

【図39】本発明の第12実施形態を示すブロック図である。

【図40】発信グループテーブル及び受信グループテーブルの一例を示す図である。

【図41】本発明の第13実施形態を示すブロック図である。

【図42】発信グループテーブル及び受信グループテーブルの一例を示す図である。

【図43】本発明の第14実施形態を示すブロック図である。

【図44】発信グループテーブル及び受信グループテーブルの一例を示す図である。

【図45】本発明の第15実施形態を示すブロック図である。

【図46】本発明の第15実施形態を示すブロック図である。

【図47】本発明の第16実施形態を示すブロック図である。

【図48】発信グループテーブル及び受信グループテーブルの一例を示す図である。

【図49】本発明の第17実施形態を示すブロック図である。

【図50】ディスプレイ上に表示されたイラストの例を示す図である。

【図51】本発明の第18実施形態を示すブロック図である。

【図52】本発明の第19実施形態を示すブロック図である。

【図53】本発明の第20実施形態を示すブロック図である。

【図54】本発明の第21実施形態を示すブロック図である。

【図55】本発明の第22実施形態を示すブロック図である。

【図56】本発明の第23実施形態を示すブロック図である。

【図57】本発明の第24実施形態を示すブロック図である。

【図58】本発明の第25実施形態を示すブロック図である。

【図59】本発明の第26実施形態を示すブロック図である。

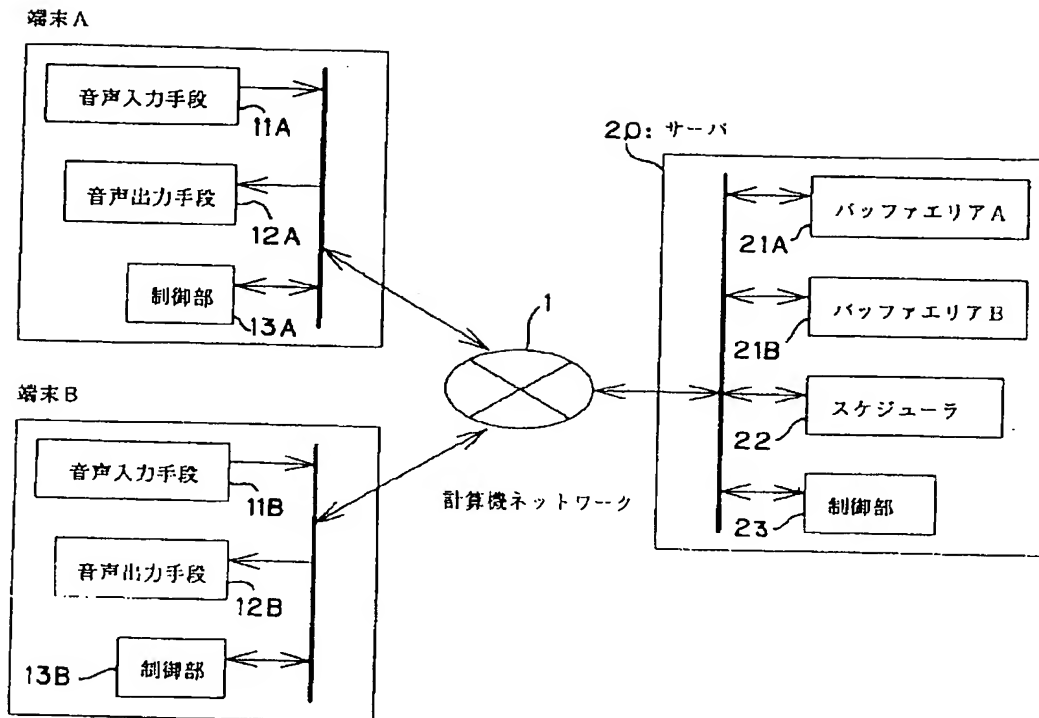
【図60】図60は、本発明の第27実施形態を示すブロック図である。

【図61】図61は、従来の音声対話システムの構成図である。

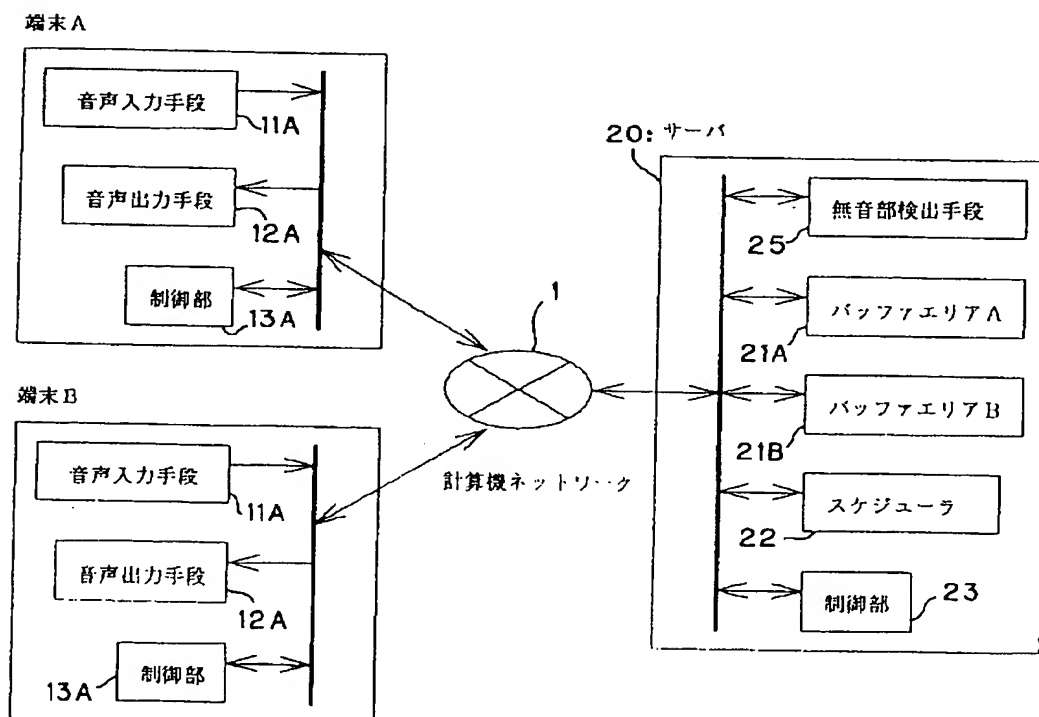
【符号の説明】

- | | |
|---------------------|----------------------|
| 1 回線 | 30 世界管理情報メモリ |
| 11A, 11B 音声入力手段 | 31 最大到達距離メモリ |
| 12A, 12B 音声出力手段 | 32 エフェクタ |
| 13A, 13B 制御部 | 33 グループ管理テーブル |
| 14A, 14B 発言要求手段 | 34 受信グループテーブル |
| 15A, 15B 発言許可表示手段 | 35 発信グループテーブル |
| 16A, 16B 発言終了通知手段 | 36 特権レベルテーブル |
| 17A, 17B アバタ管理情報メモリ | 37 セキュリティ検査手段 |
| 18A, 18B 世界管理情報メモリ | 38 エフェクタ |
| 19A, 19B 発信モード選択手段 | 39 効果バッファ |
| 20 サーバ | 40 効果音管理情報メモリ |
| 21A, 21B バッファメモリ | 41 反復間隔メモリ |
| 22 スケジューラ | 42 音声認識手段 |
| 23 制御部 | 43 音声合成手段 |
| 24 同時発言者数管理手段 | 44 翻訳手段 |
| 25 無音部検出手段 | 45 使用言語テーブル |
| 26 内部時計 | 46 検索手段 |
| 27 内部時間補正值計算部 | 47 語句テーブル |
| 28 出力バッファエリア | 50 ハードディスク |
| 29 ミキサ | 51 検索手段 |
| | 52 話速変換手段 |
| | 101B 聴取モード選択手段 |
| | 102A, 102B パスワード入力手段 |
| | 103A, 103B 描画手段 |
| | 104A, 104B 座標入力手段 |
| | 105A, 105B 表示手段 |
| | 106A, 106B モード設定手段 |
| | 107B 音声認識手段 |

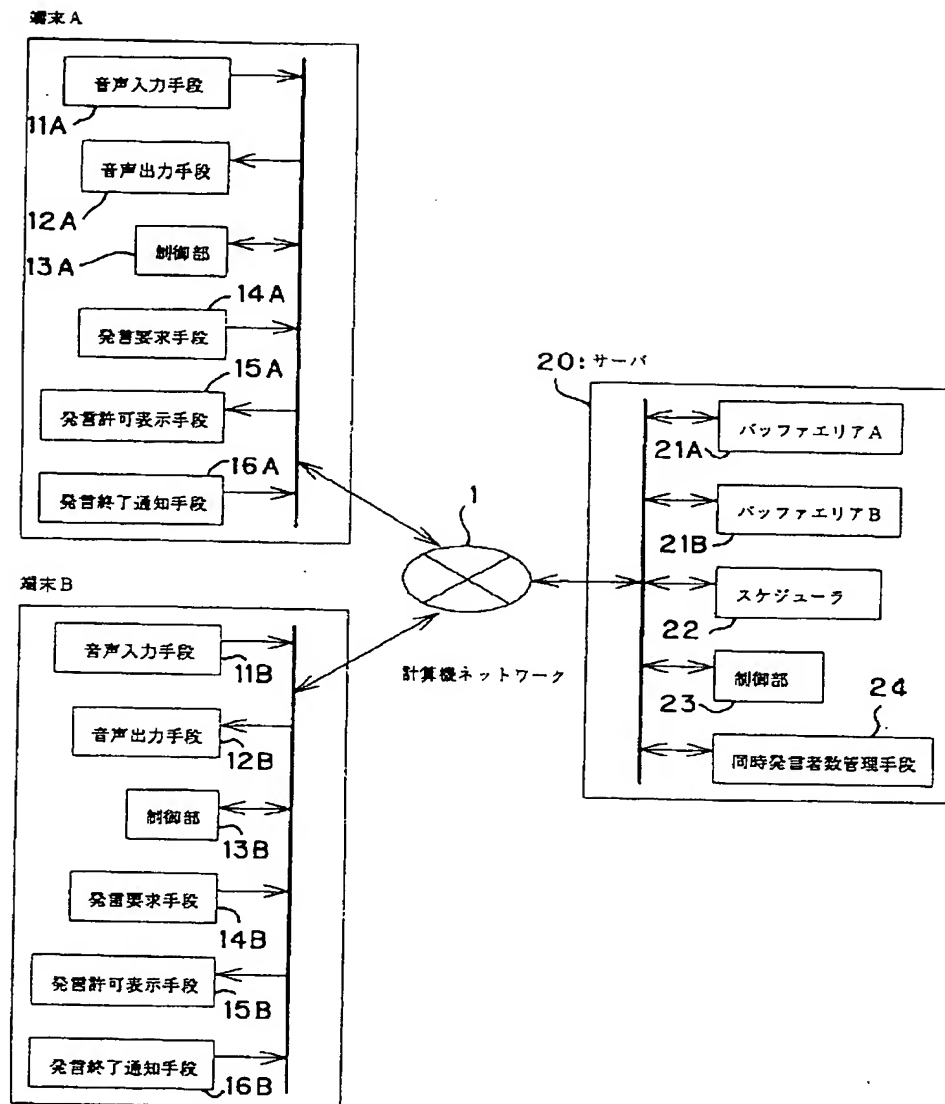
【図1】



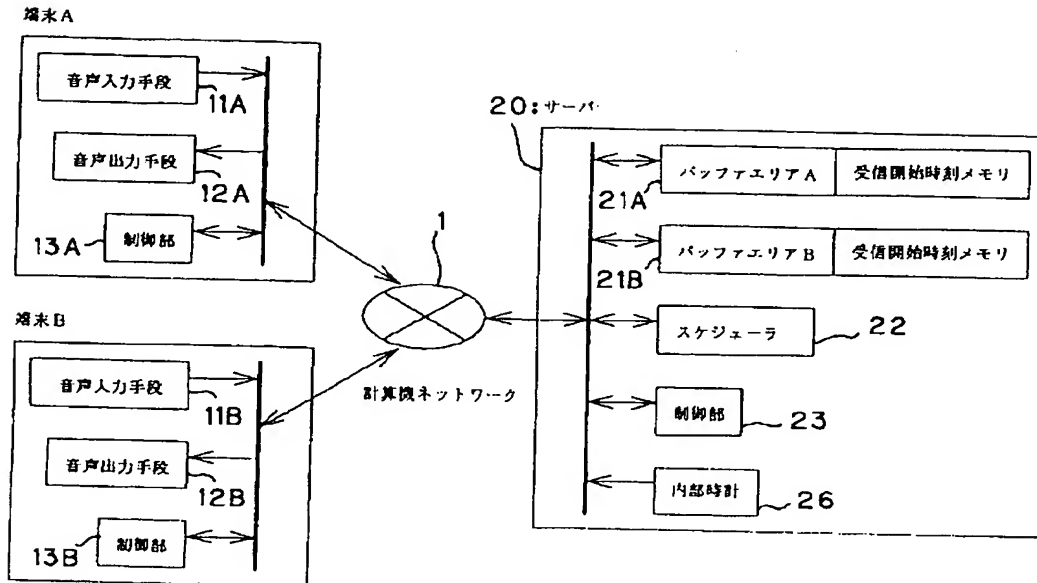
【図3】



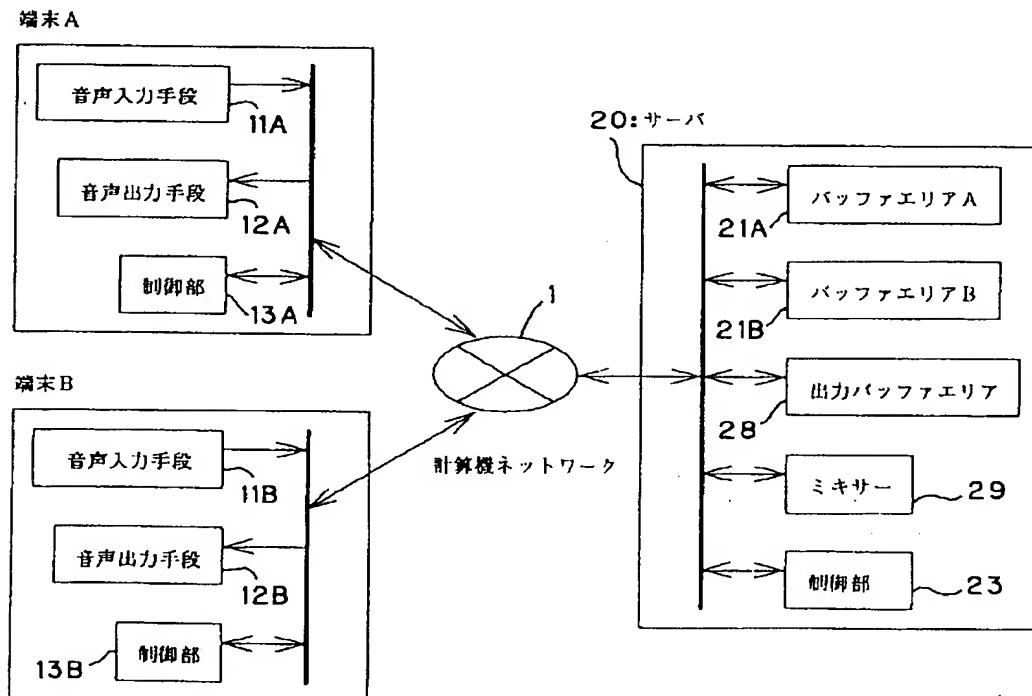
【図2】



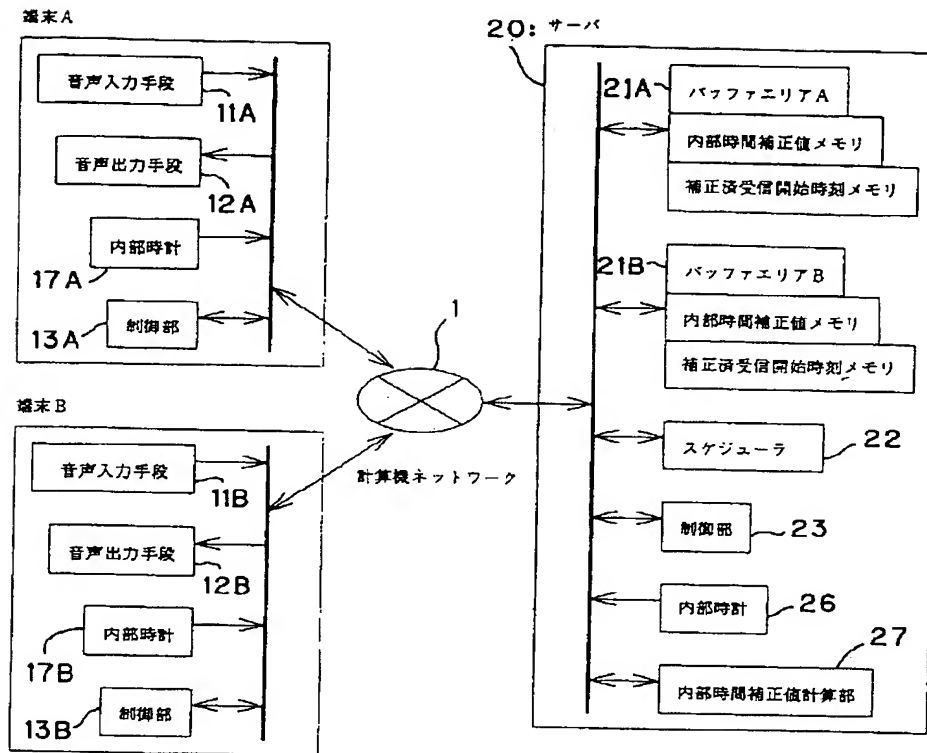
【図4】



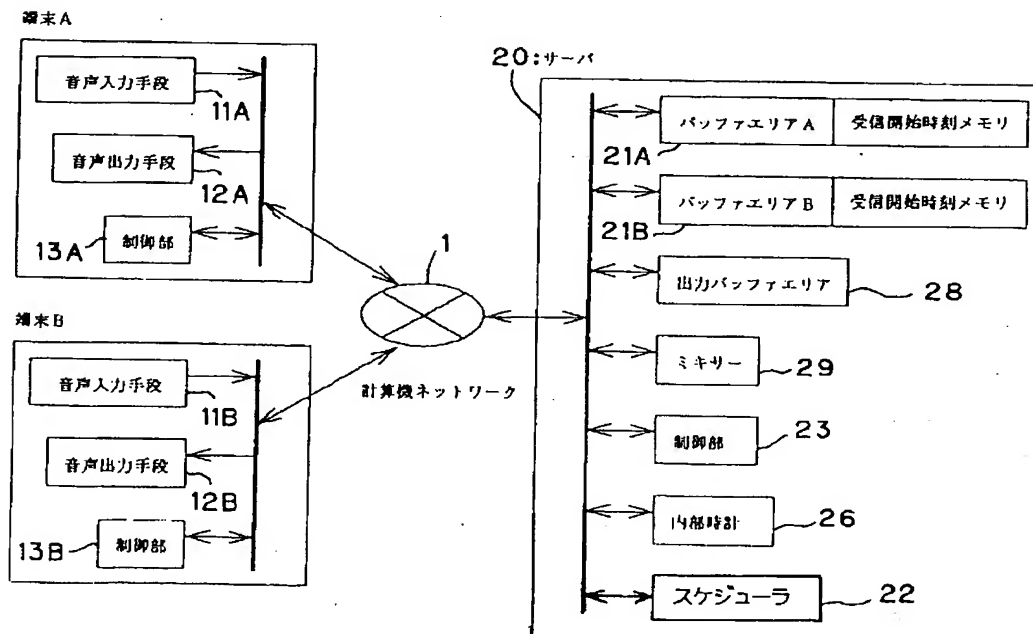
【図6】



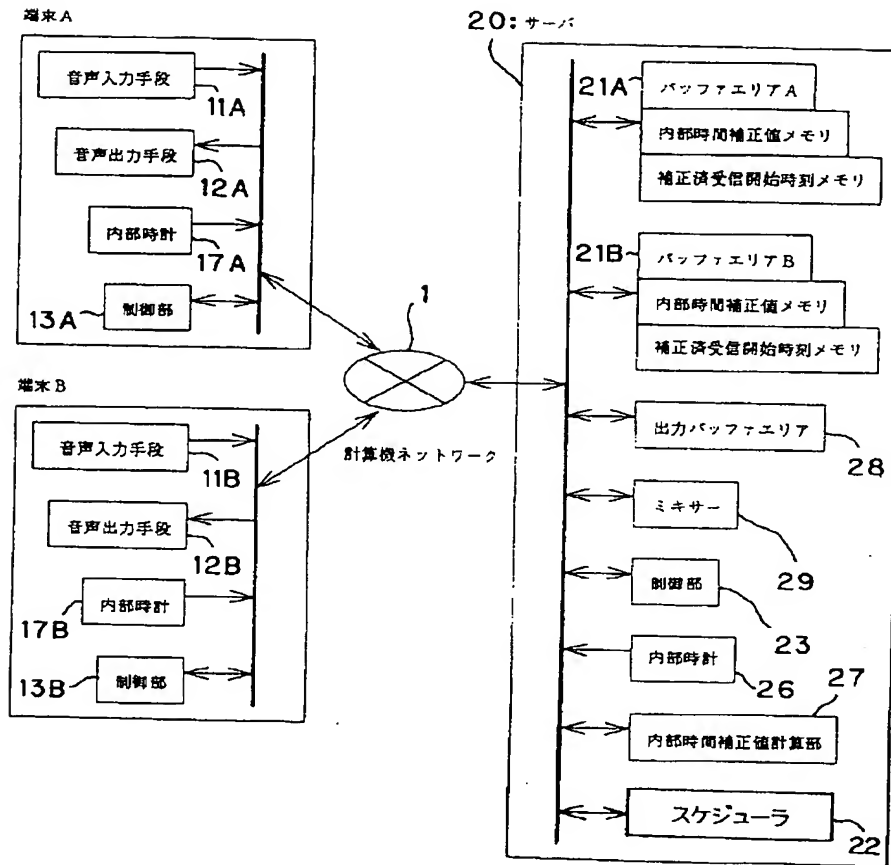
【図5】



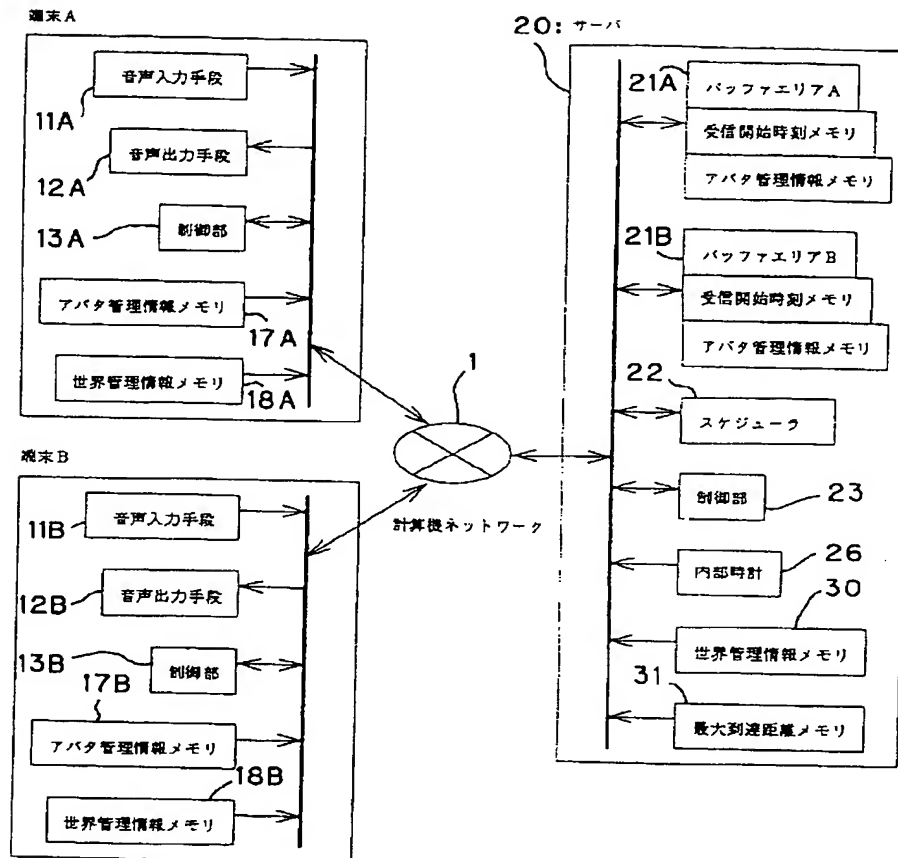
【図7】



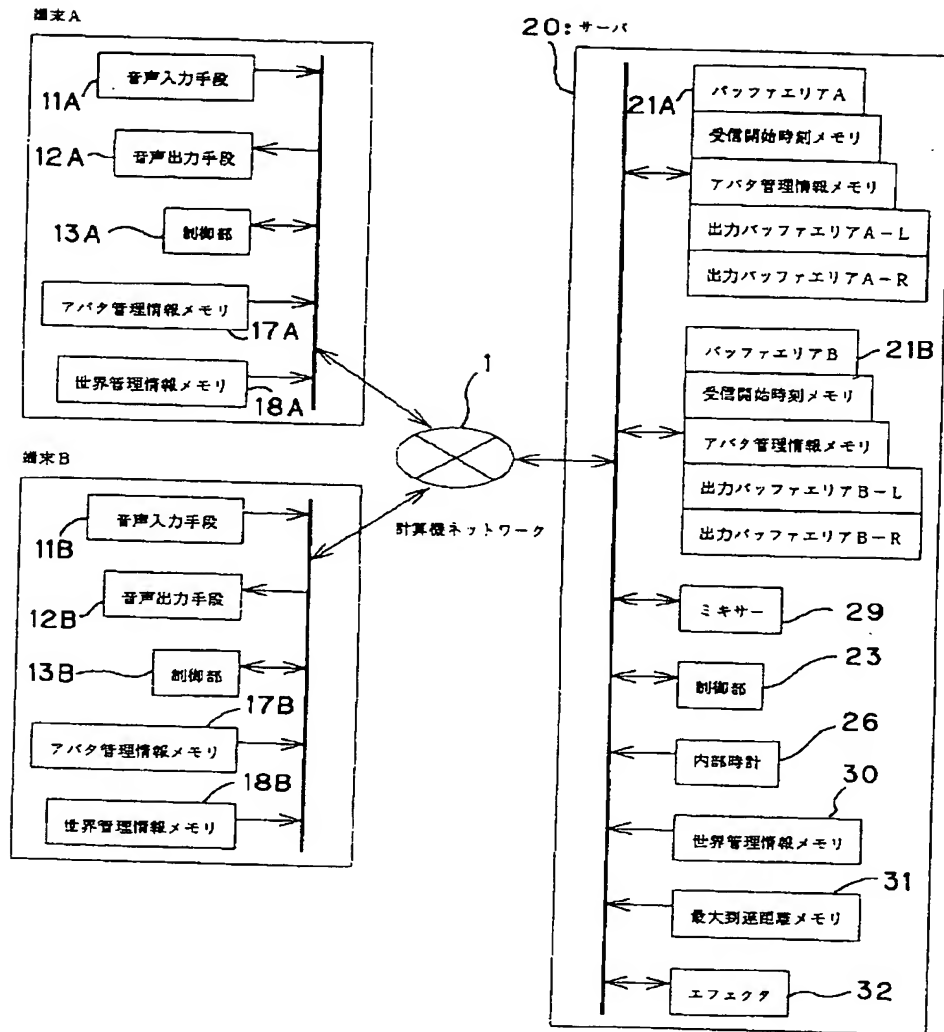
【図8】



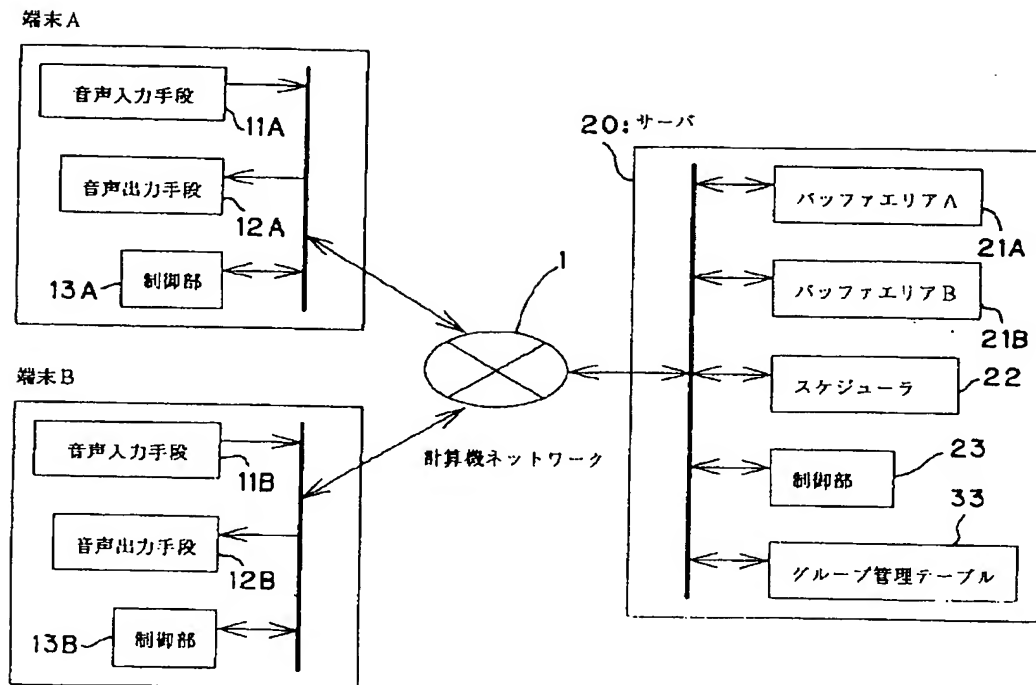
【図9】



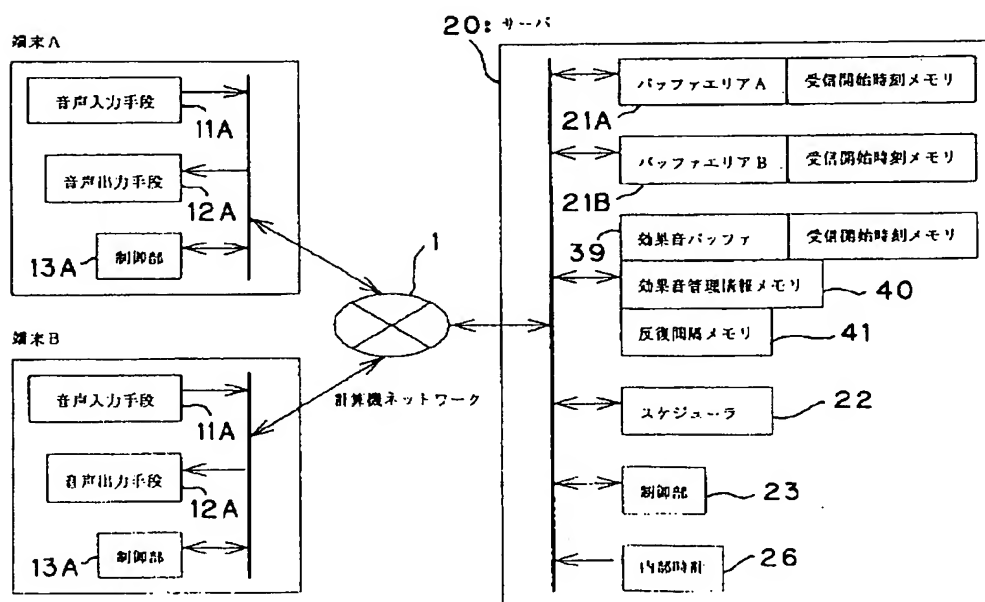
【図10】



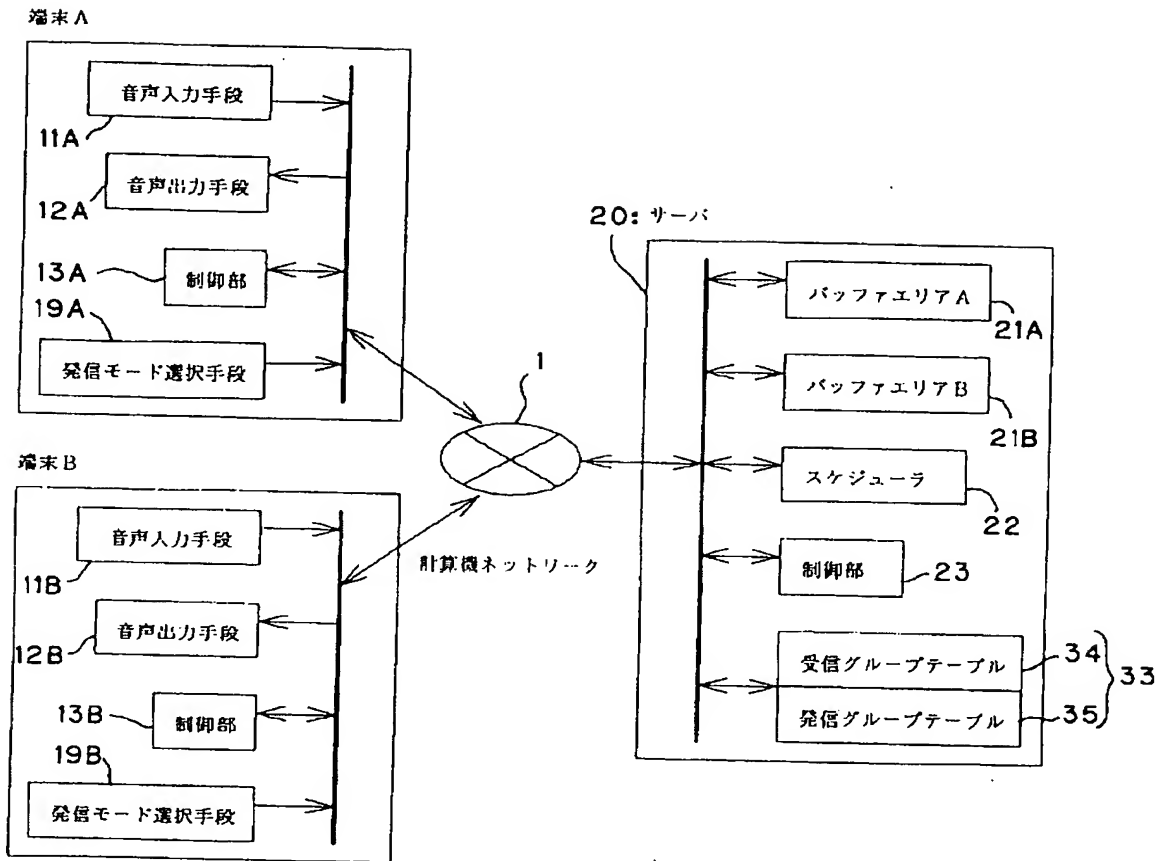
【図11】



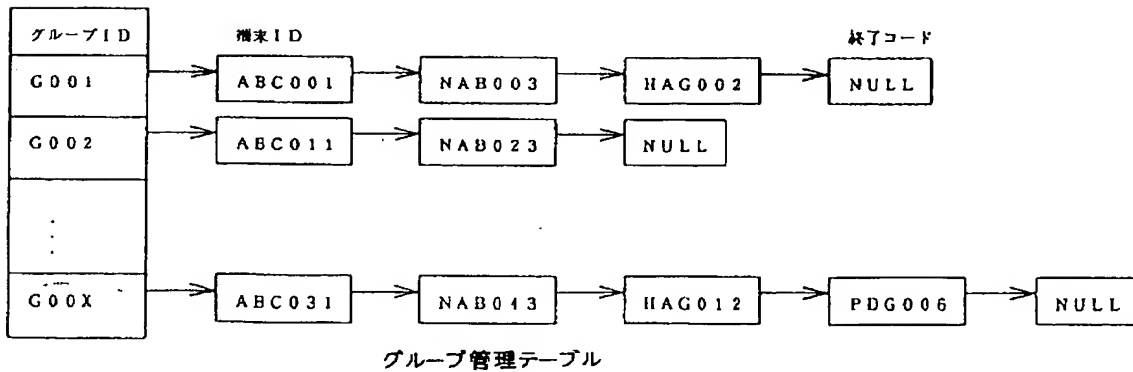
【図18】



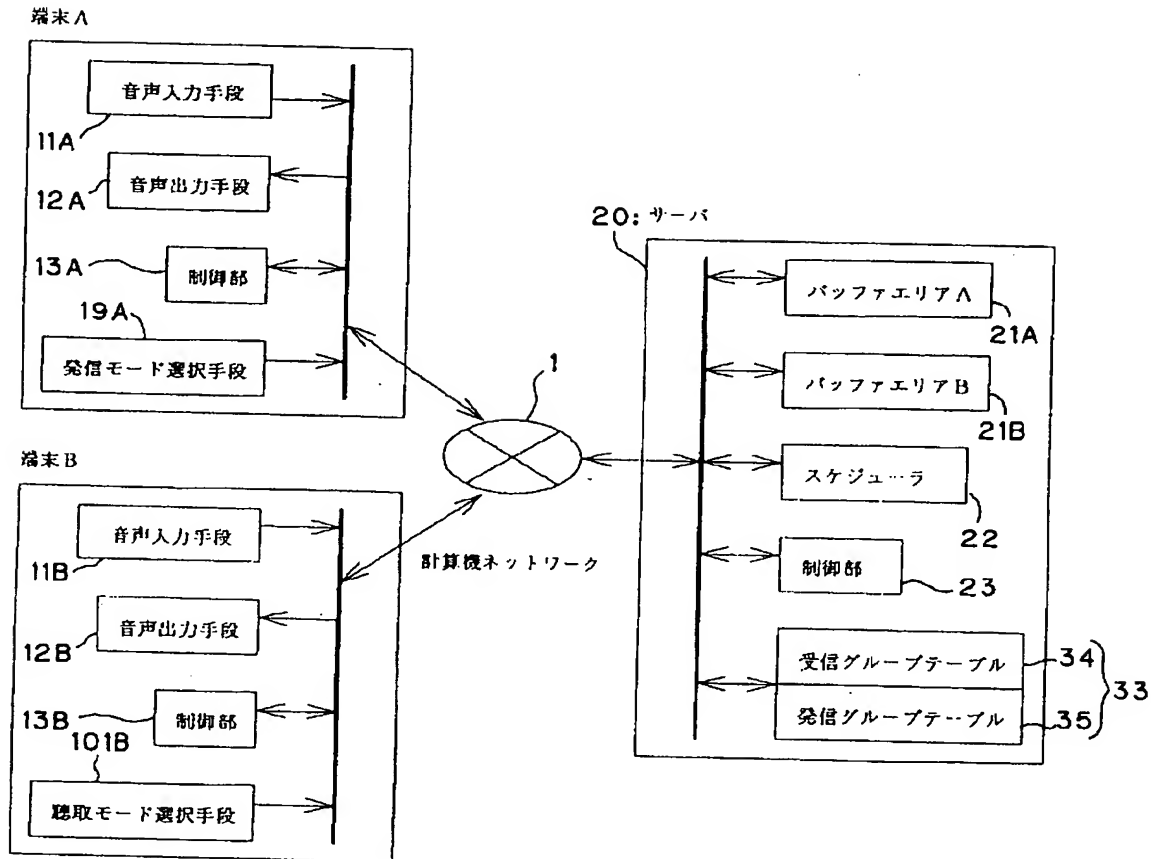
【図12】



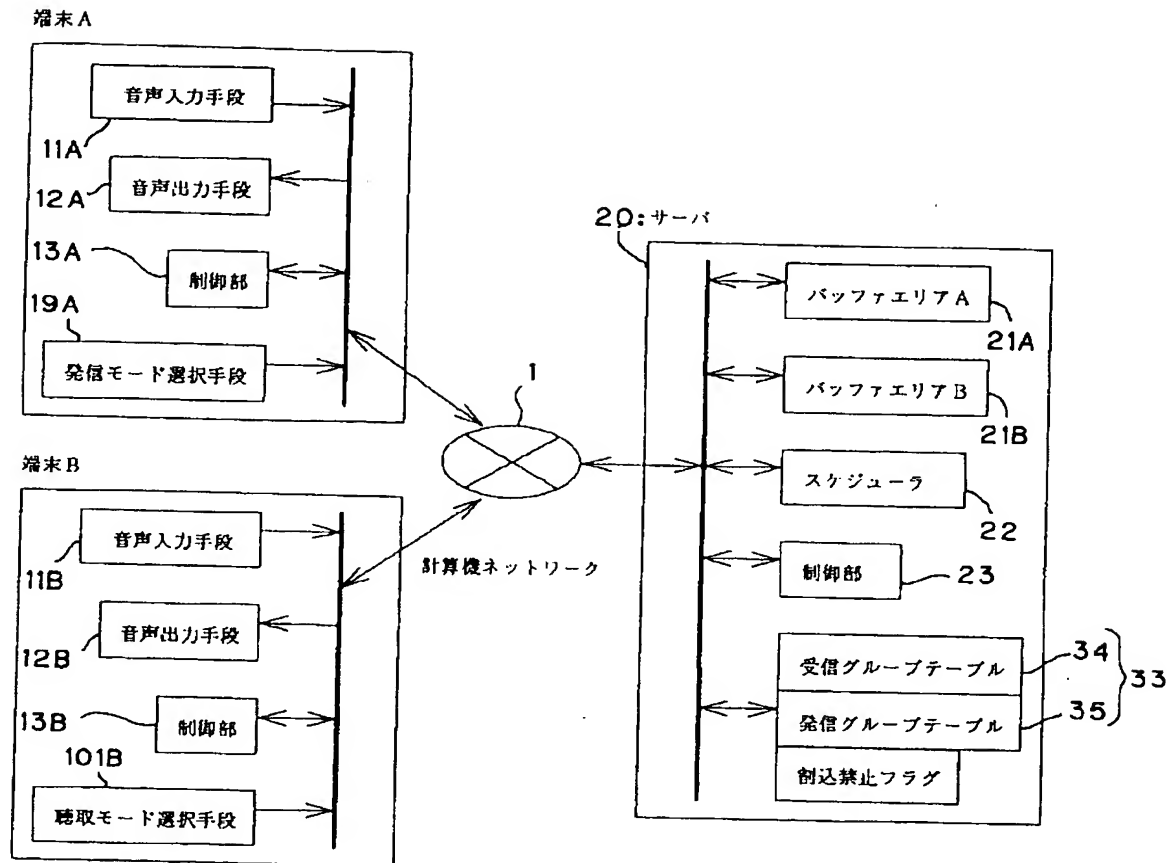
【図38】



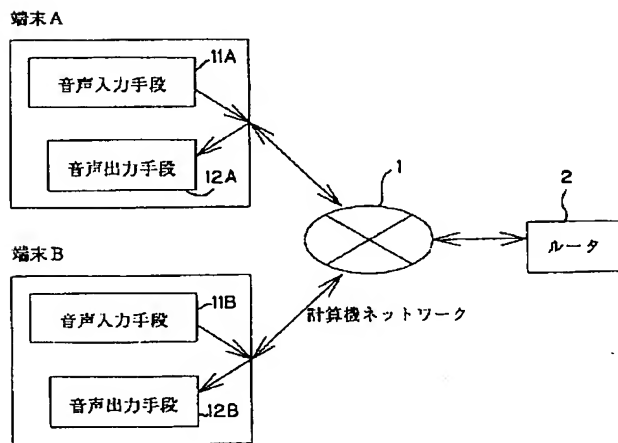
【図13】



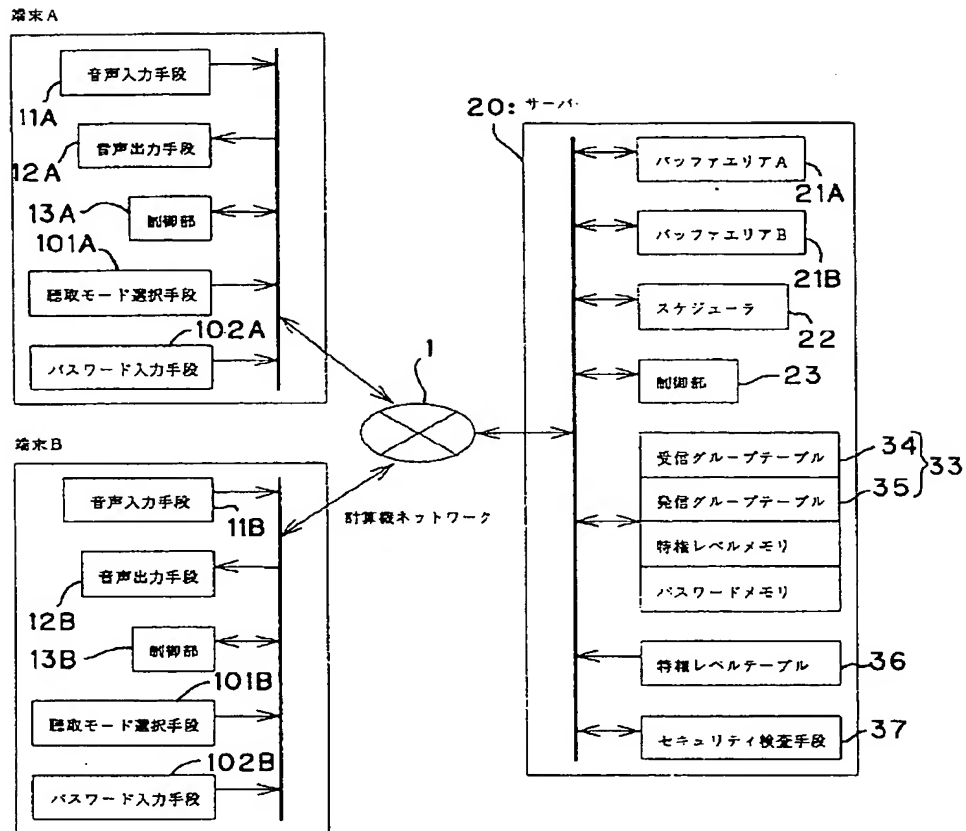
【図14】



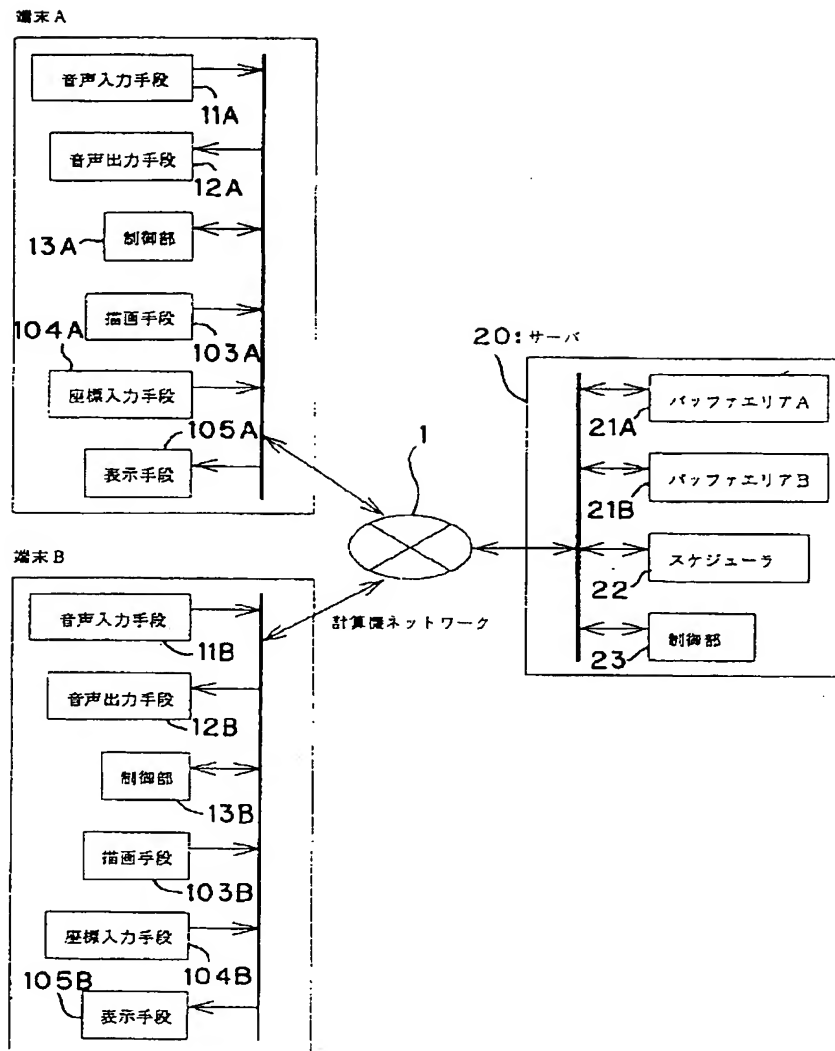
【図61】



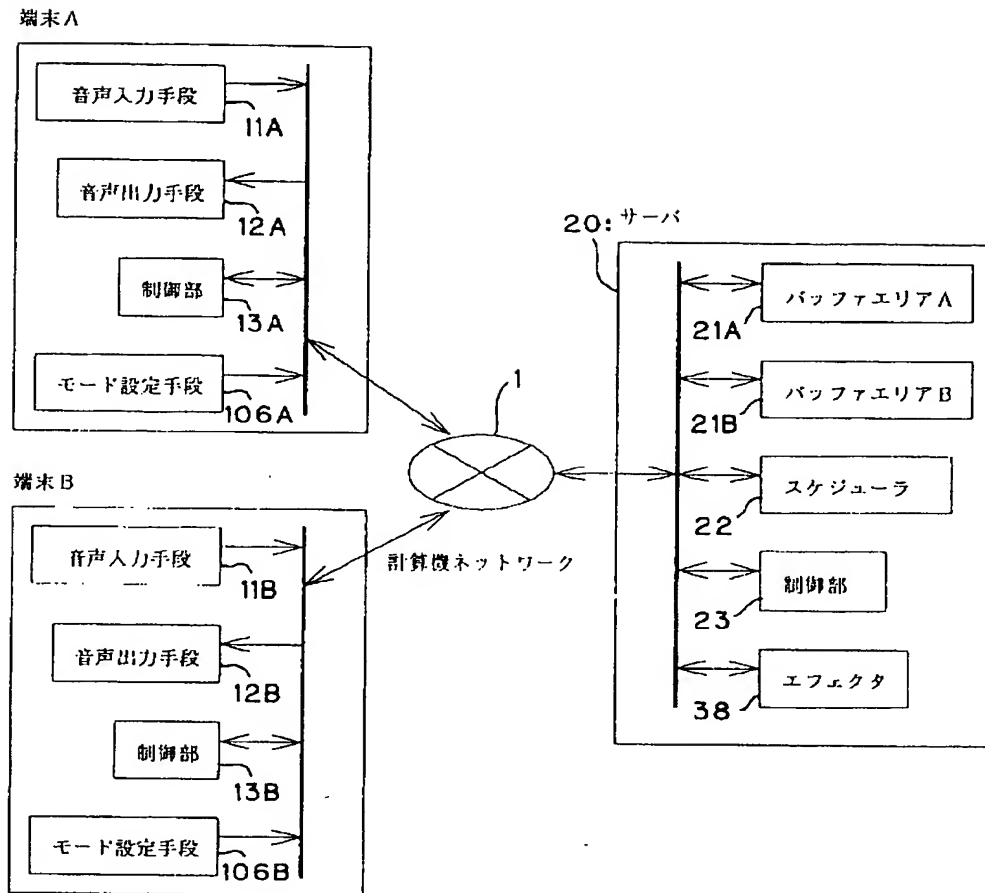
【図15】



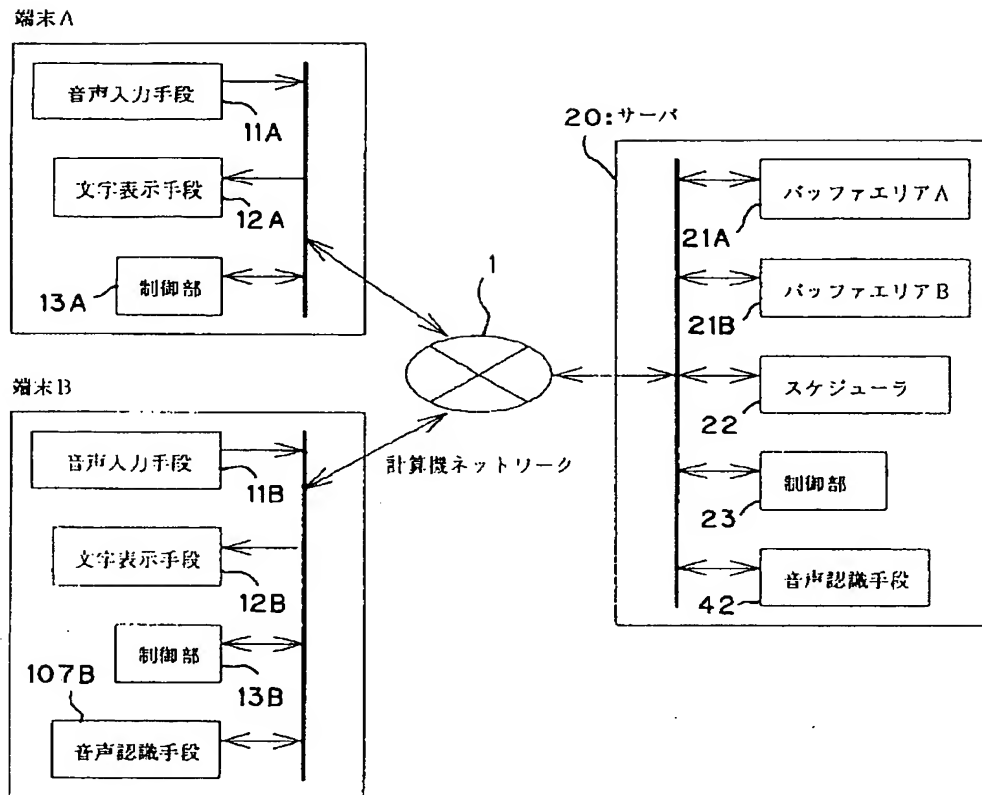
【図16】



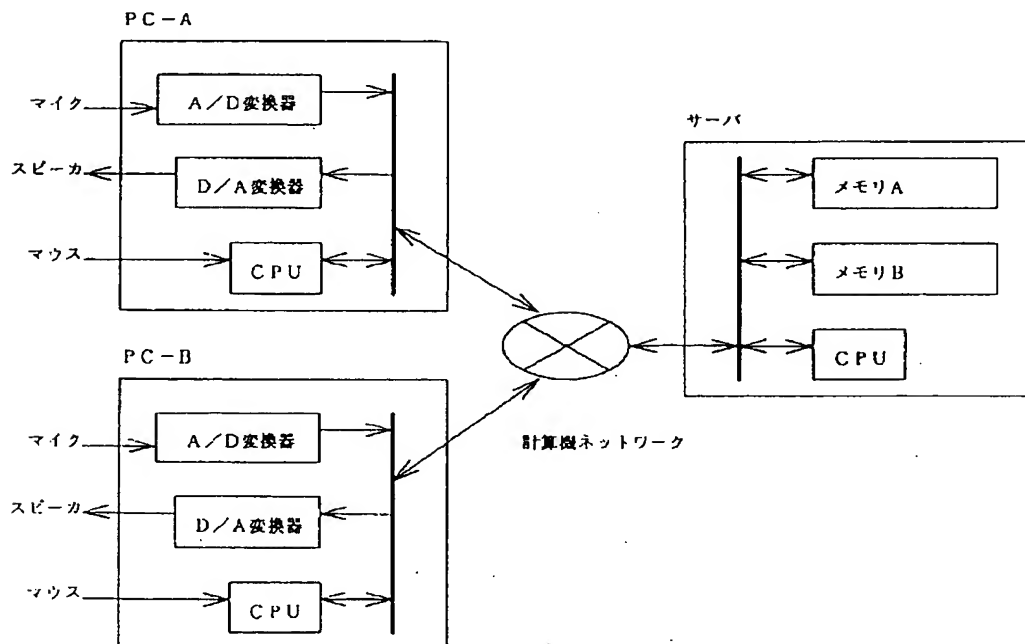
【図17】



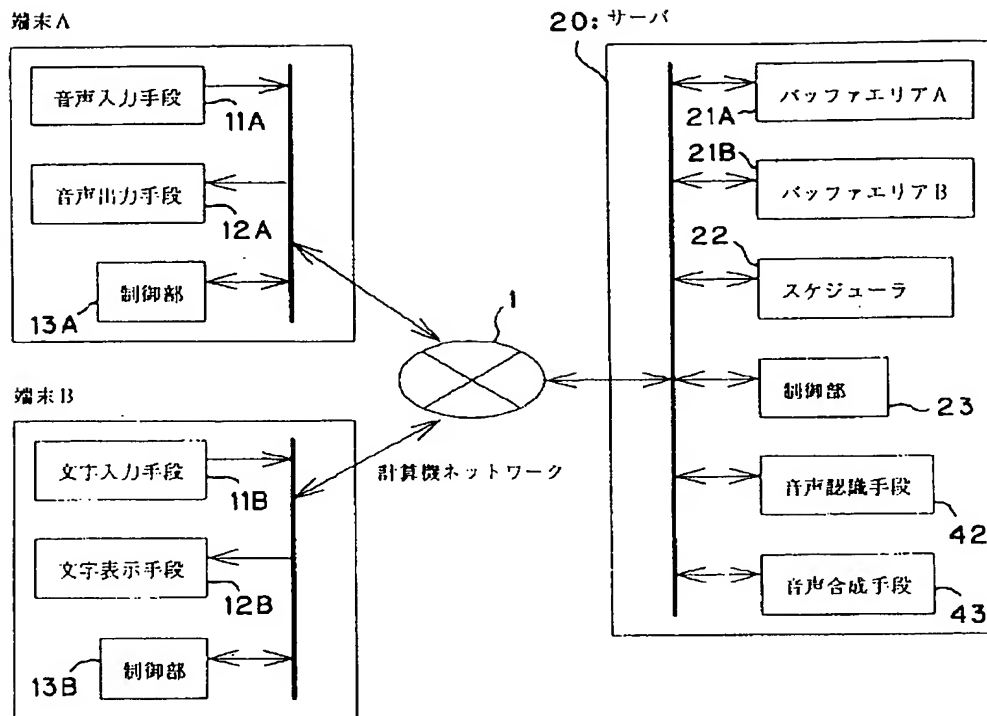
【図19】



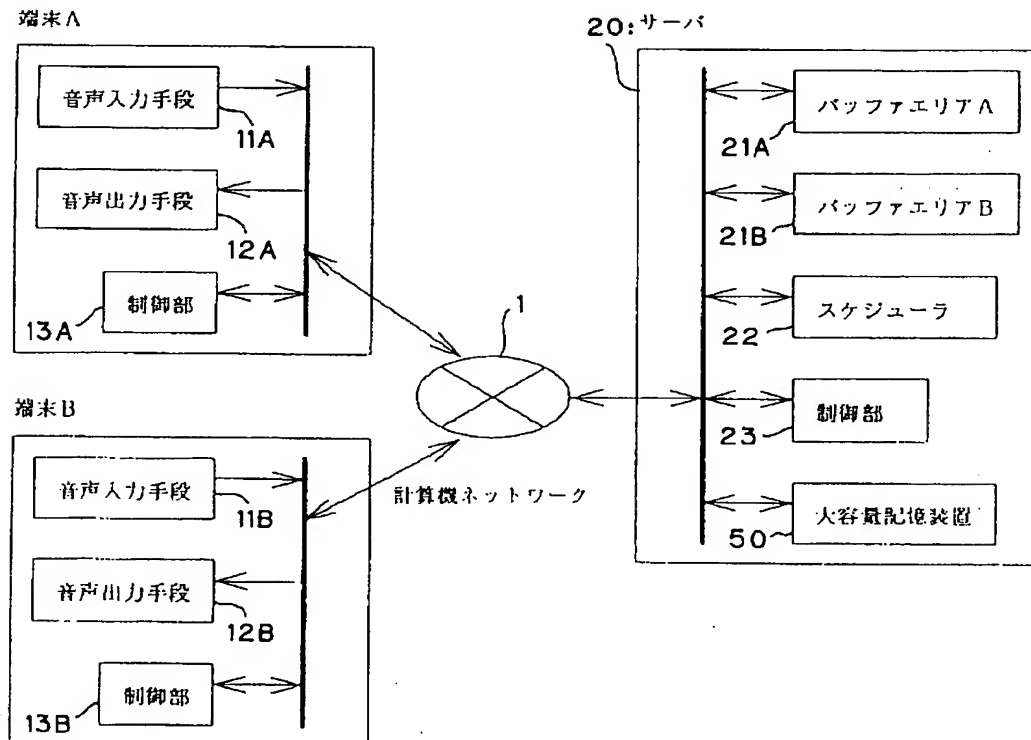
【図27】



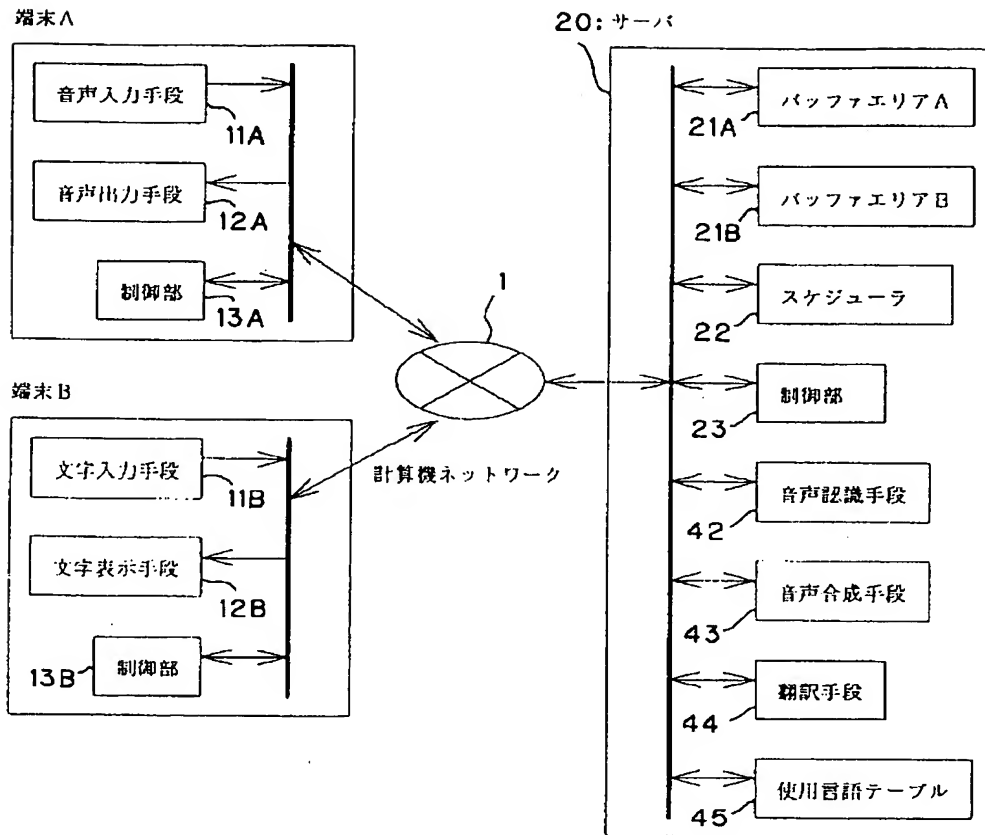
【図20】



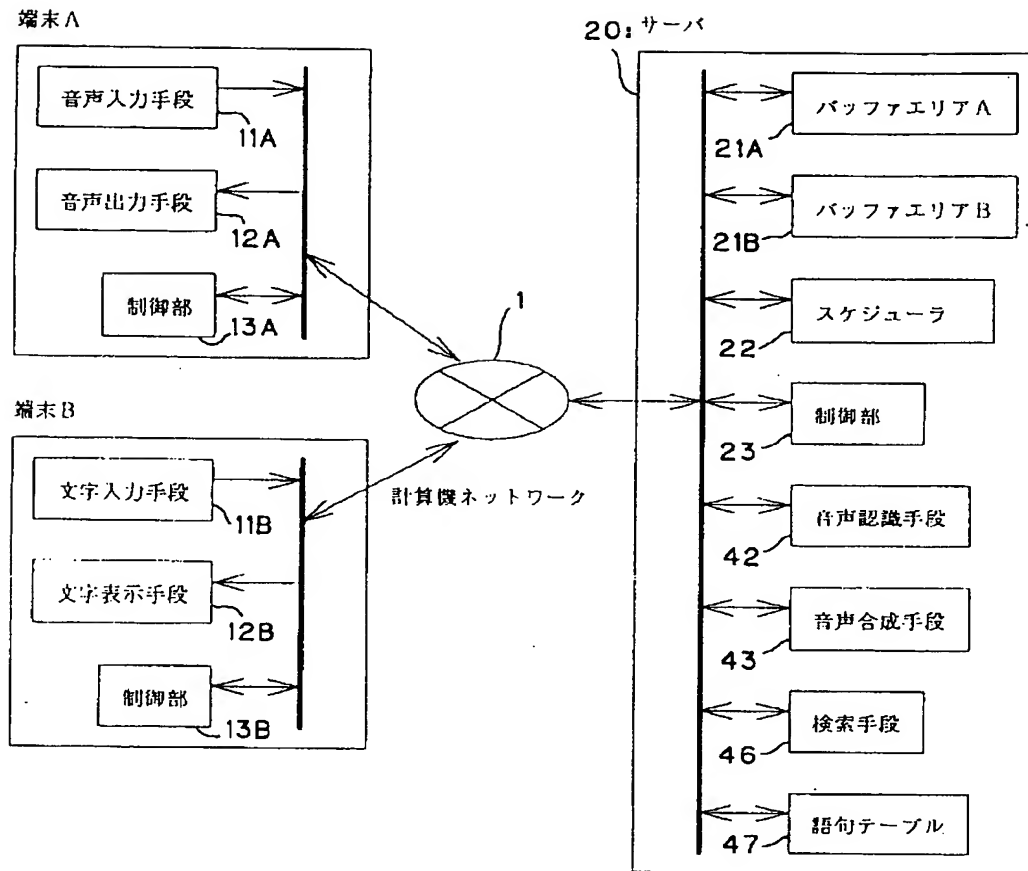
【図24】



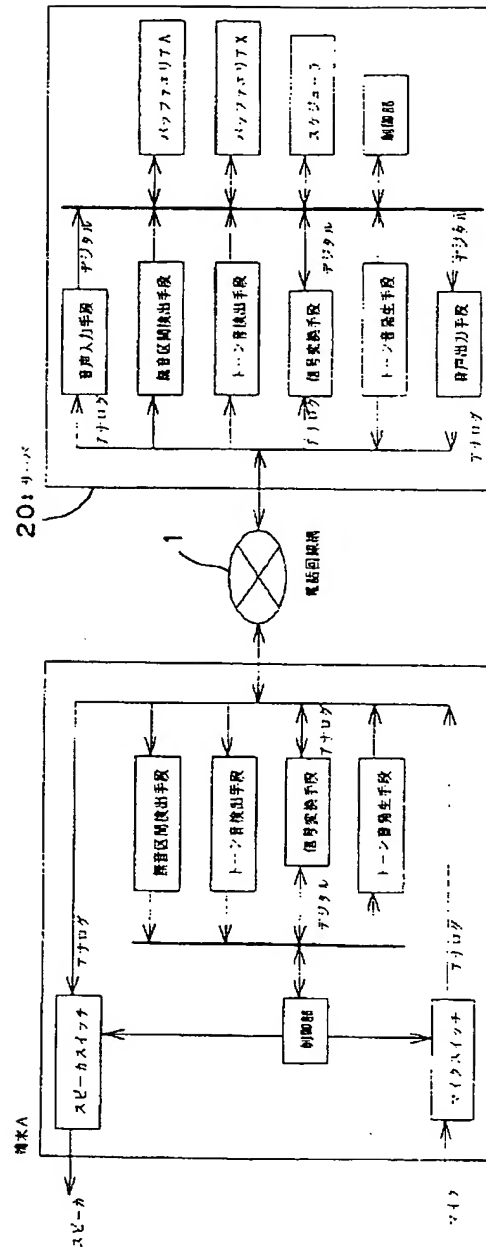
【図21】



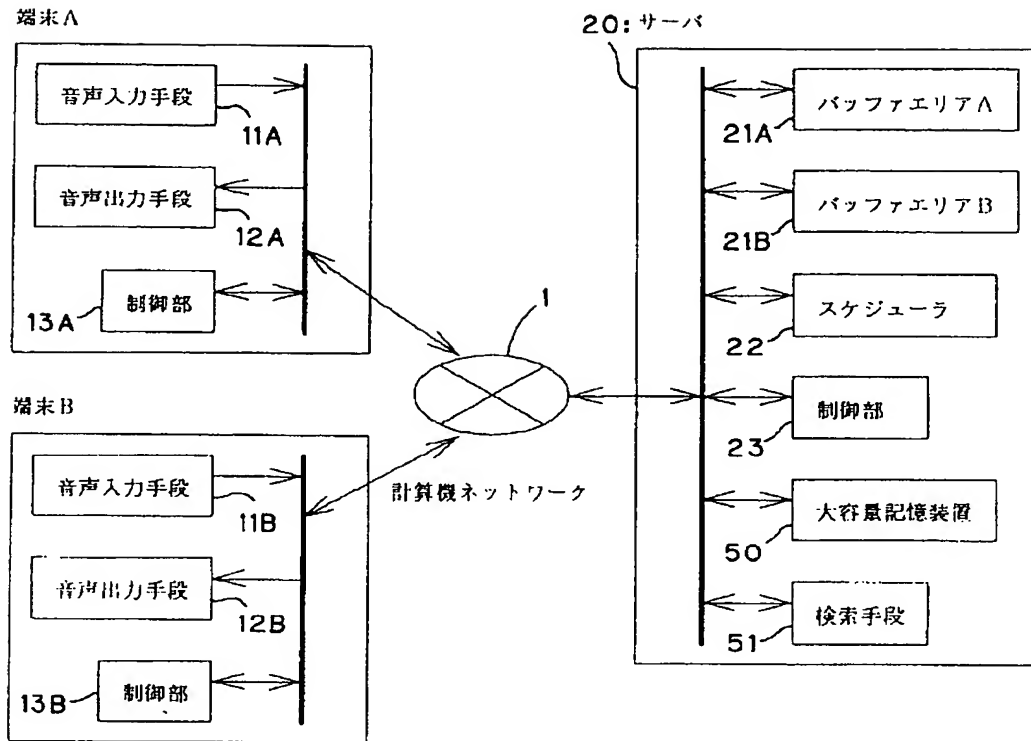
【図22】



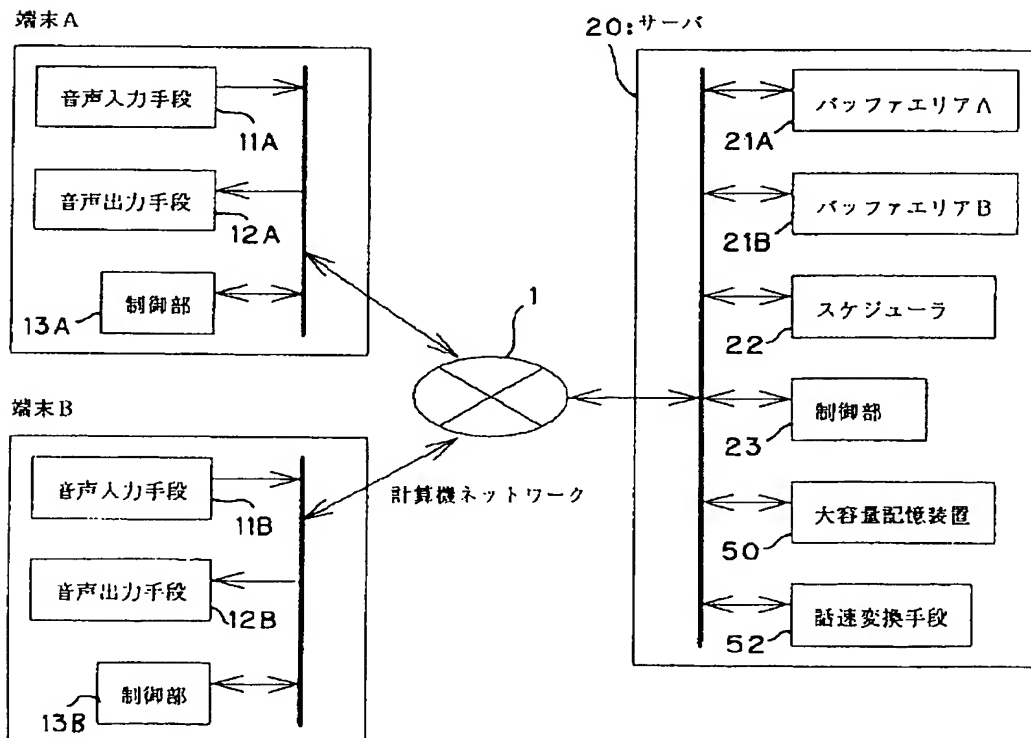
【図23】



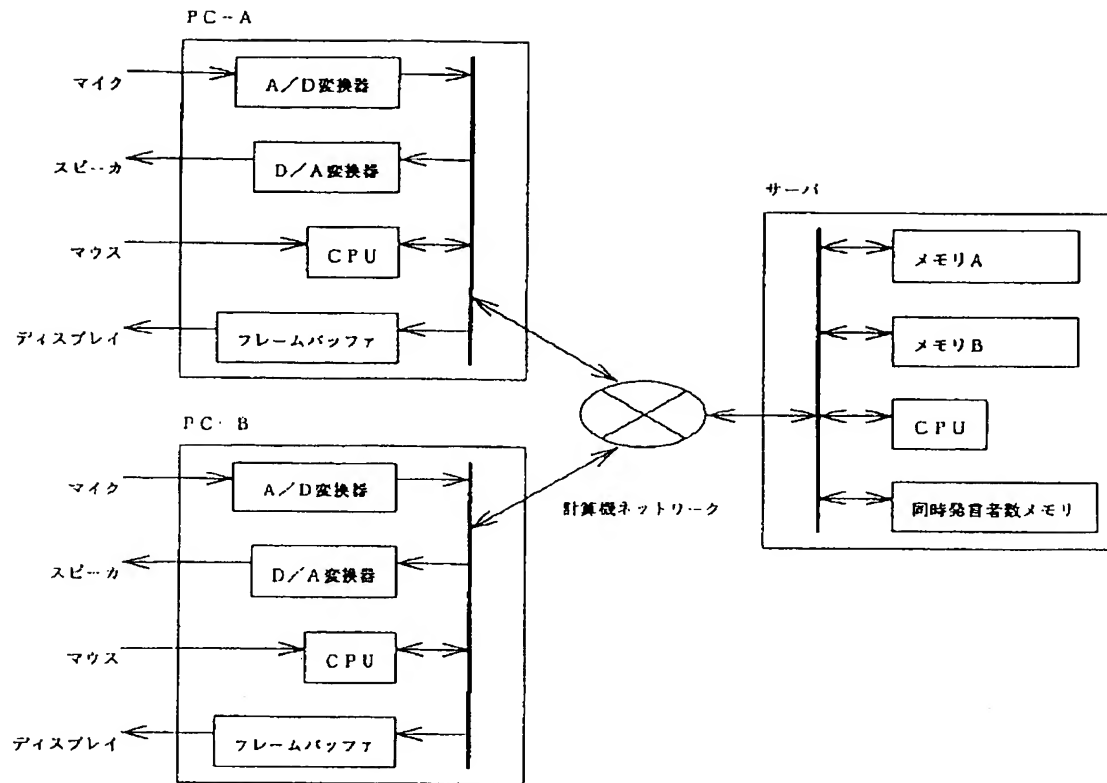
【図25】



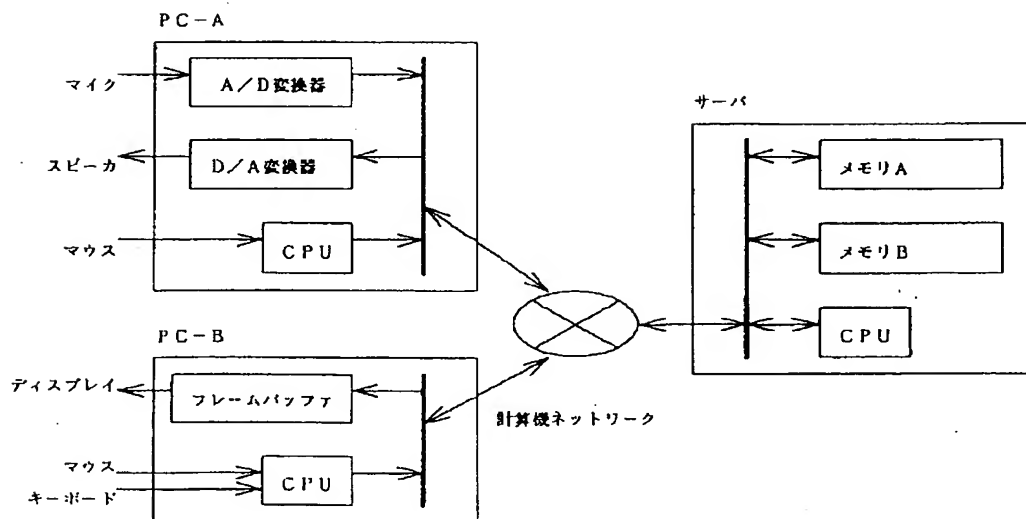
【図26】



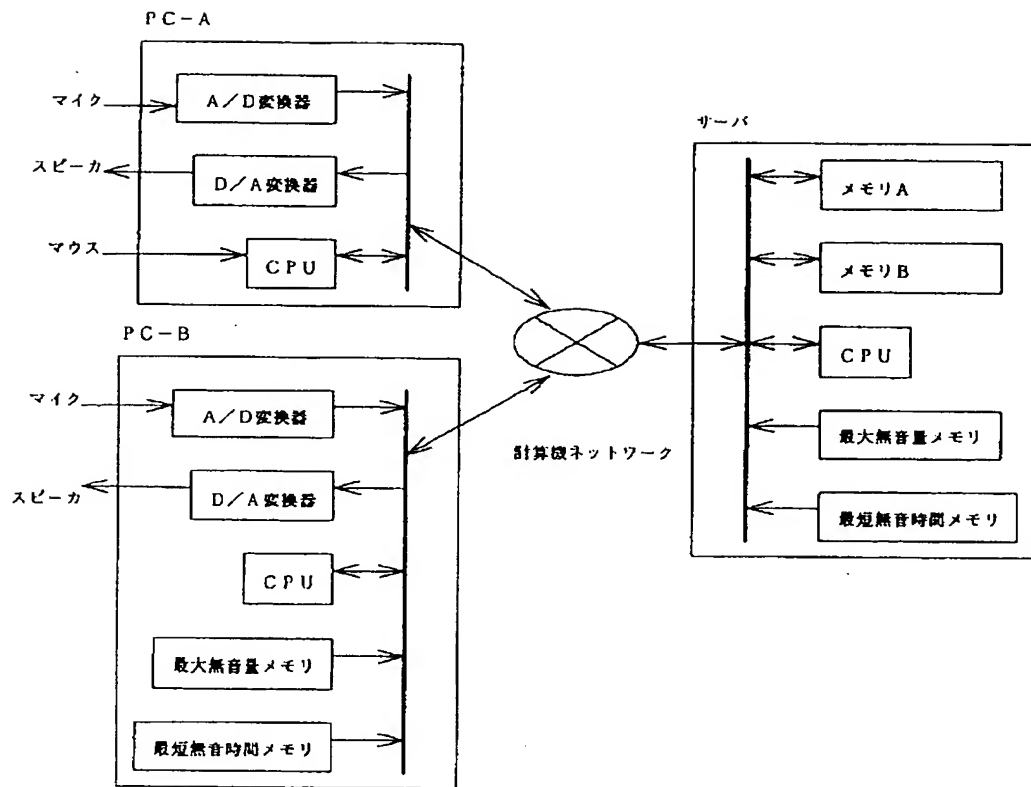
【図28】



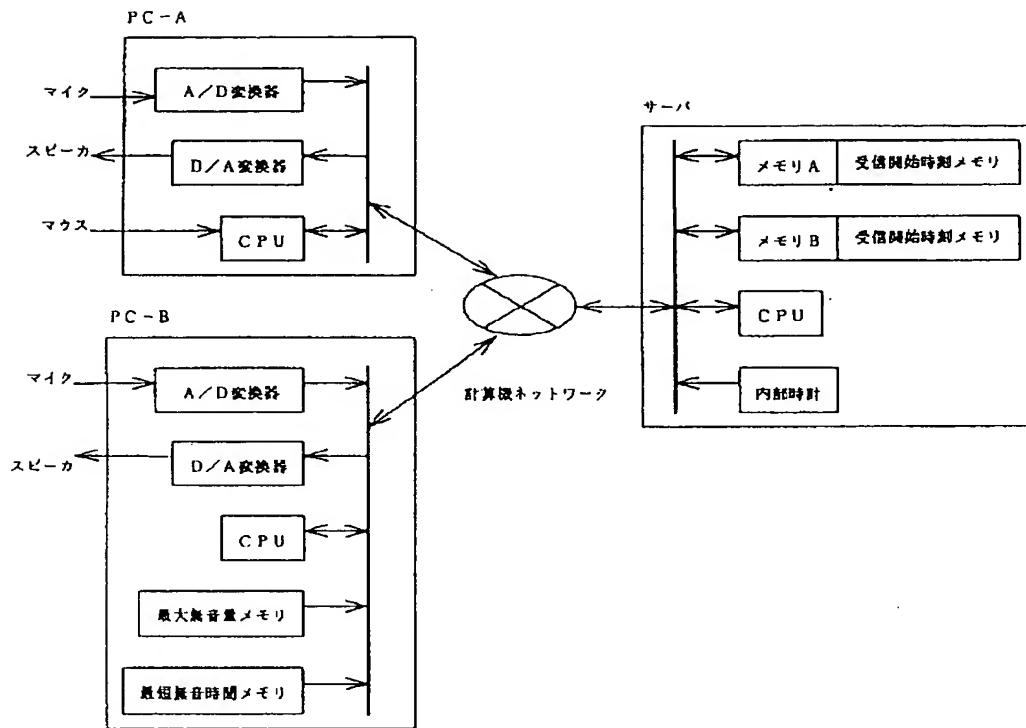
【図54】



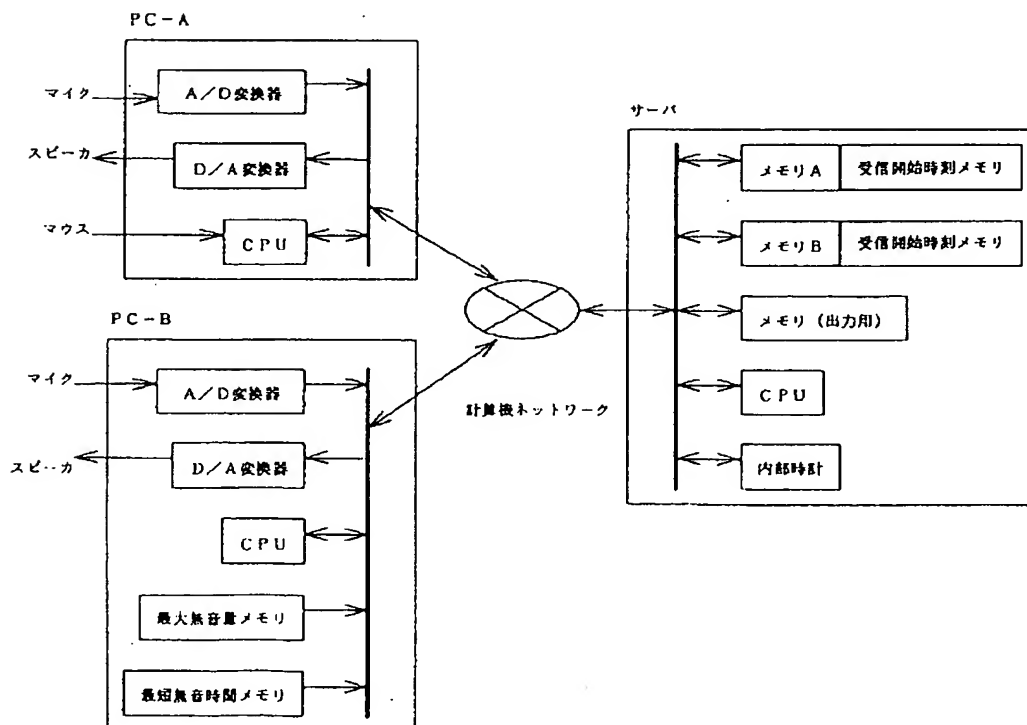
【図29】



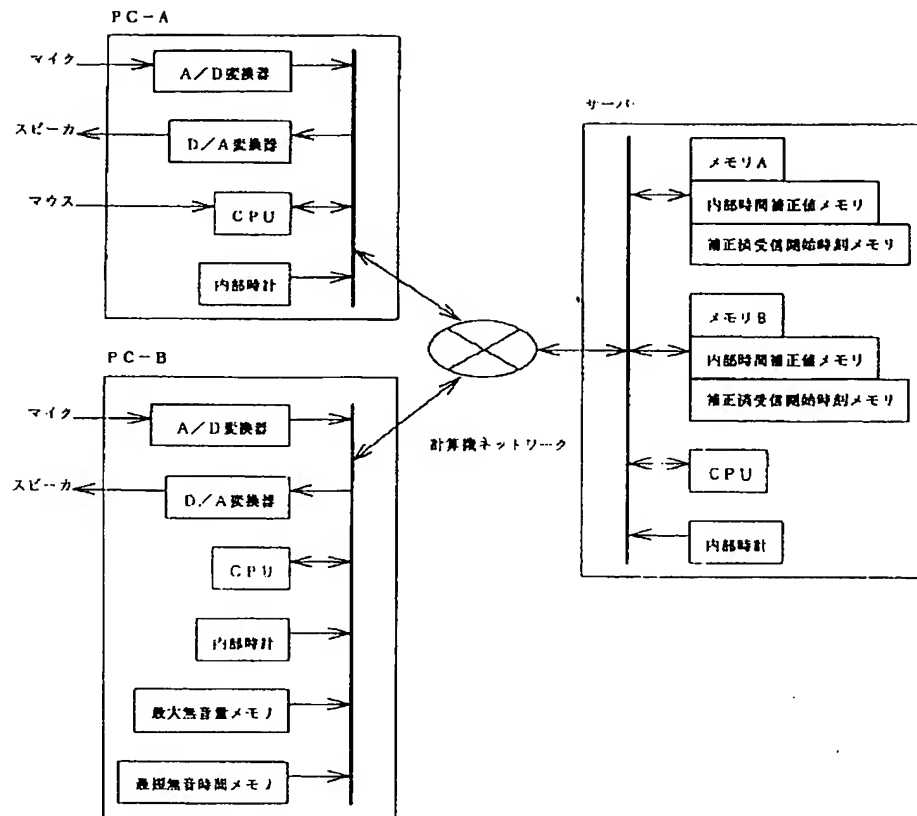
【図30】



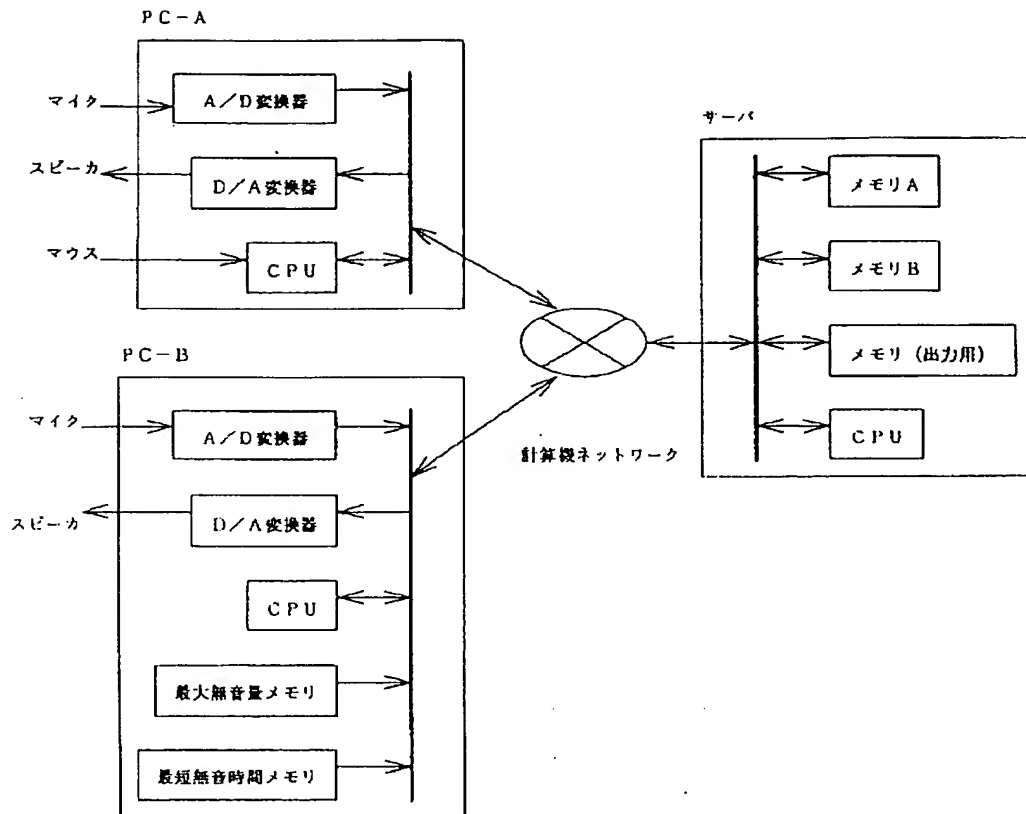
【図33】



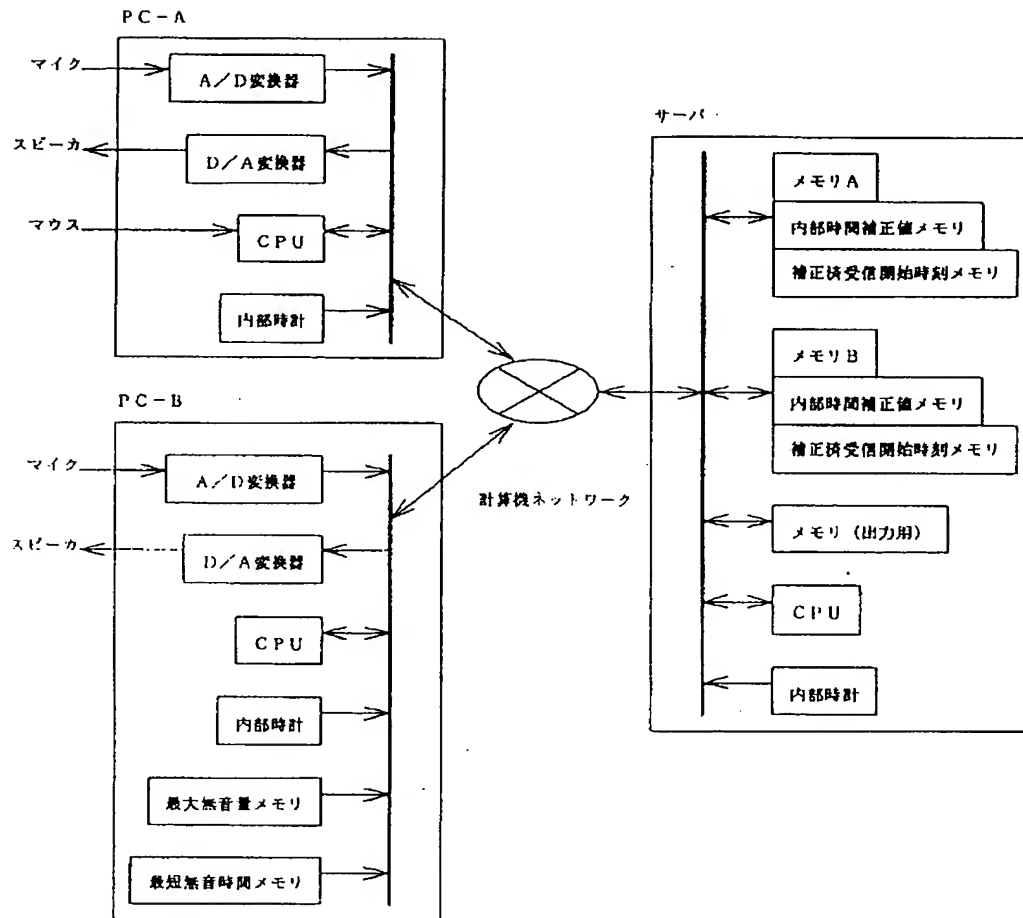
【図31】



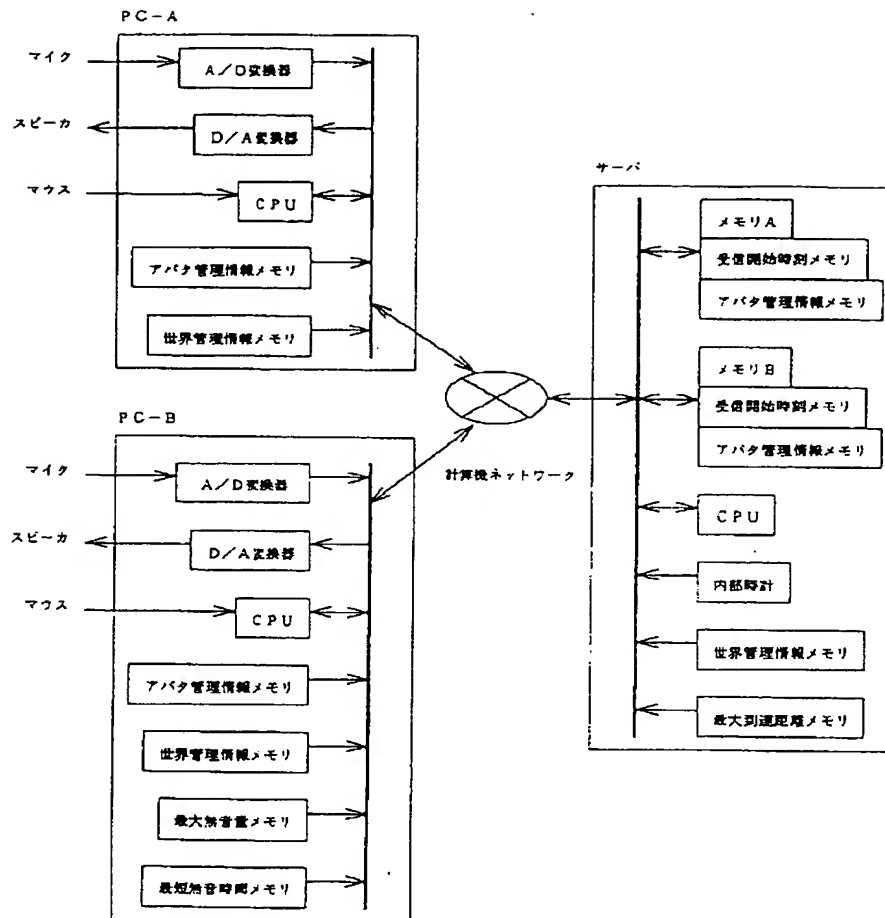
【図32】



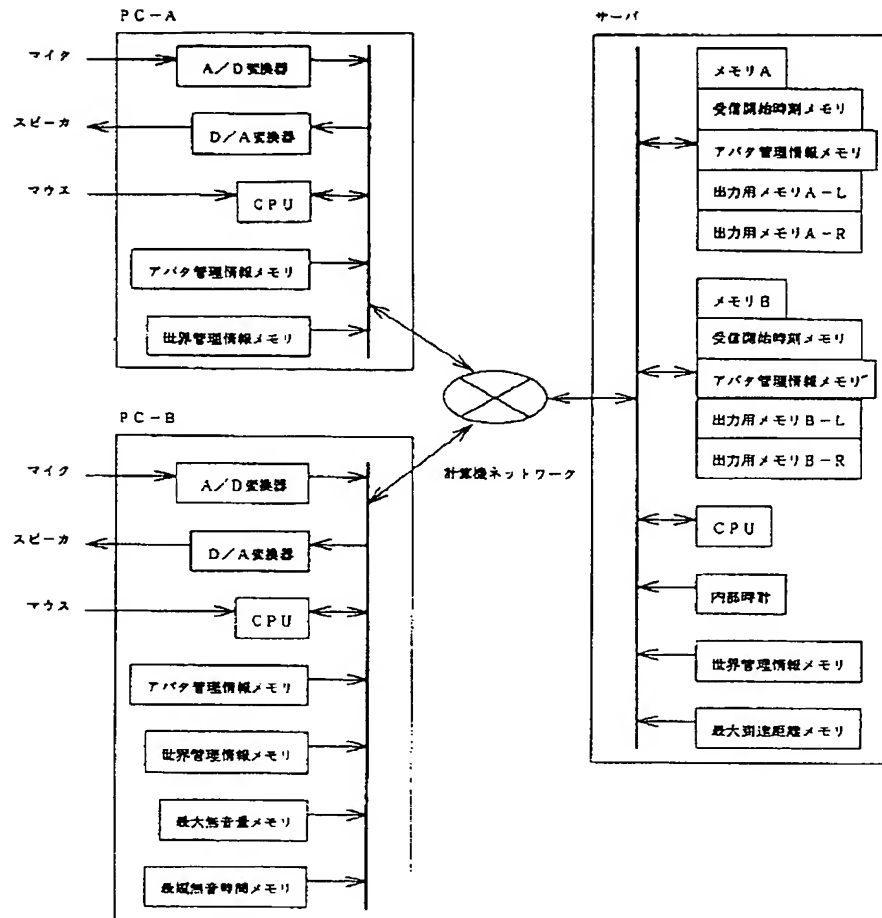
【図34】



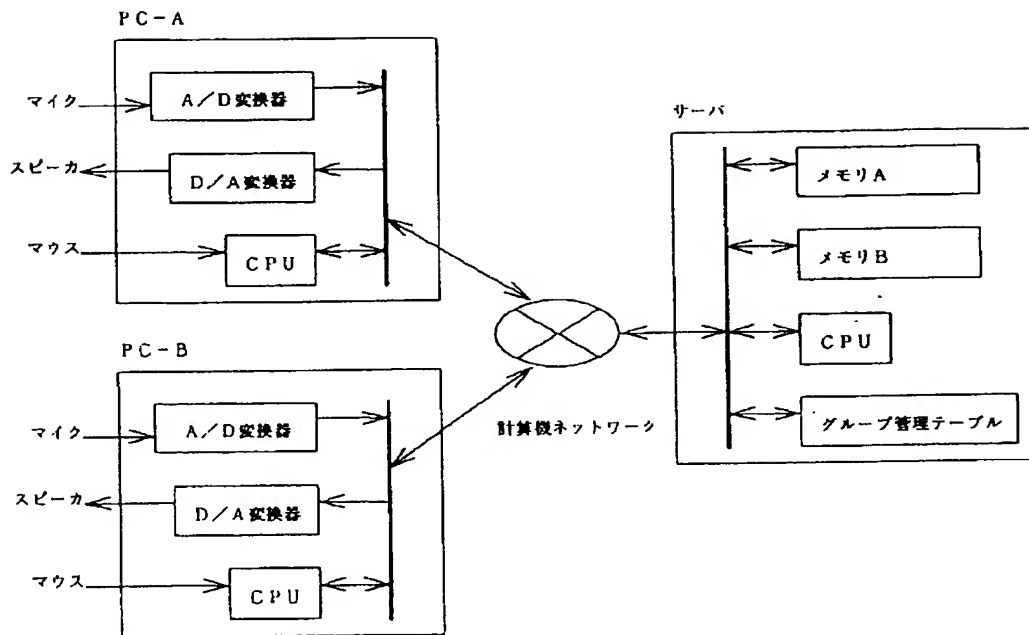
【図35】



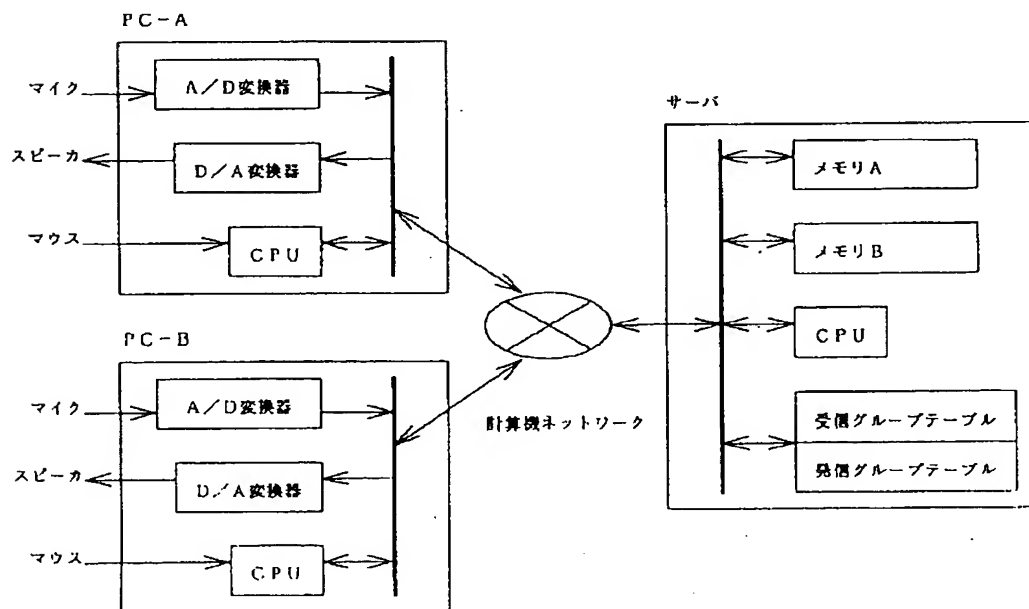
【図36】



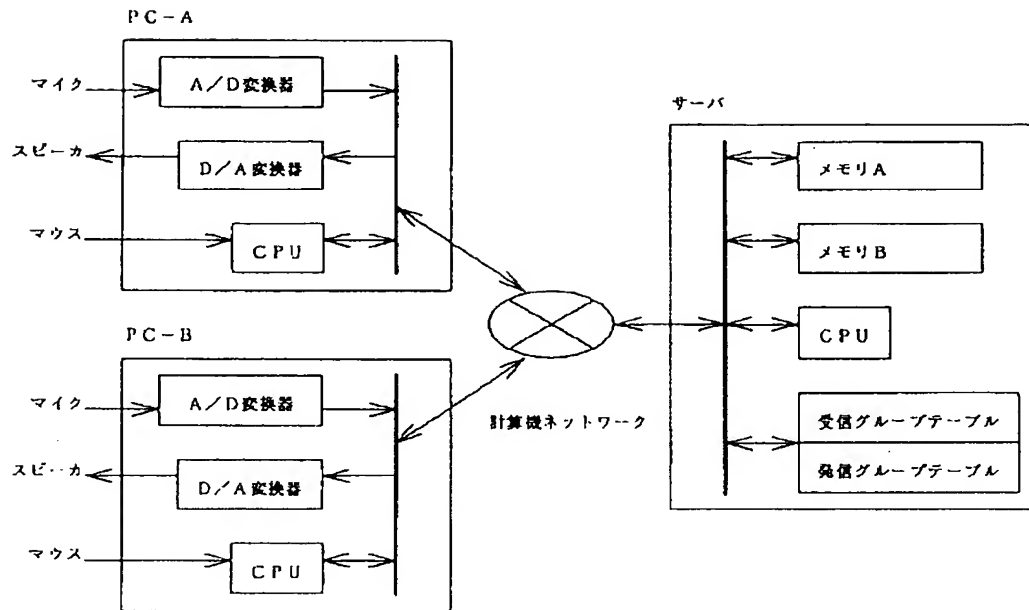
【図37】



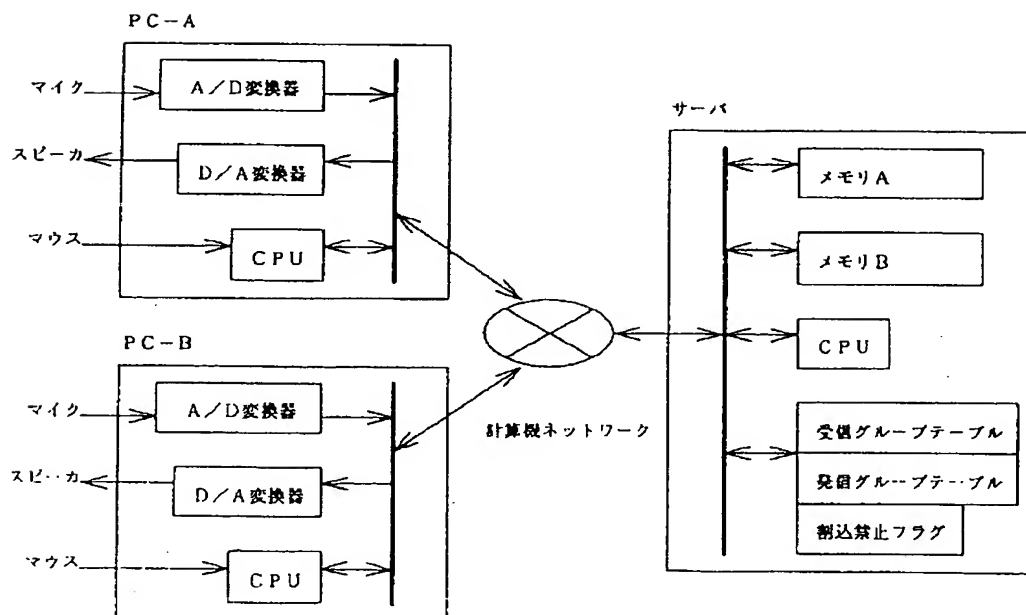
【図41】



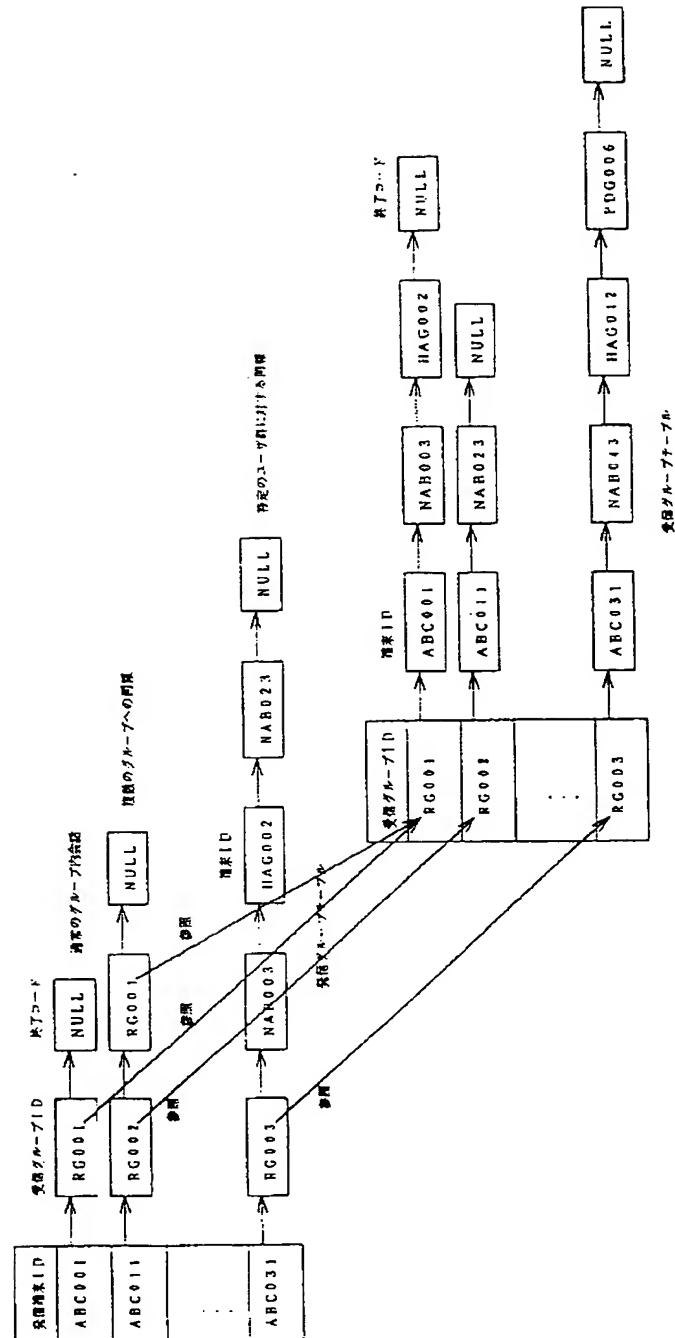
【図39】



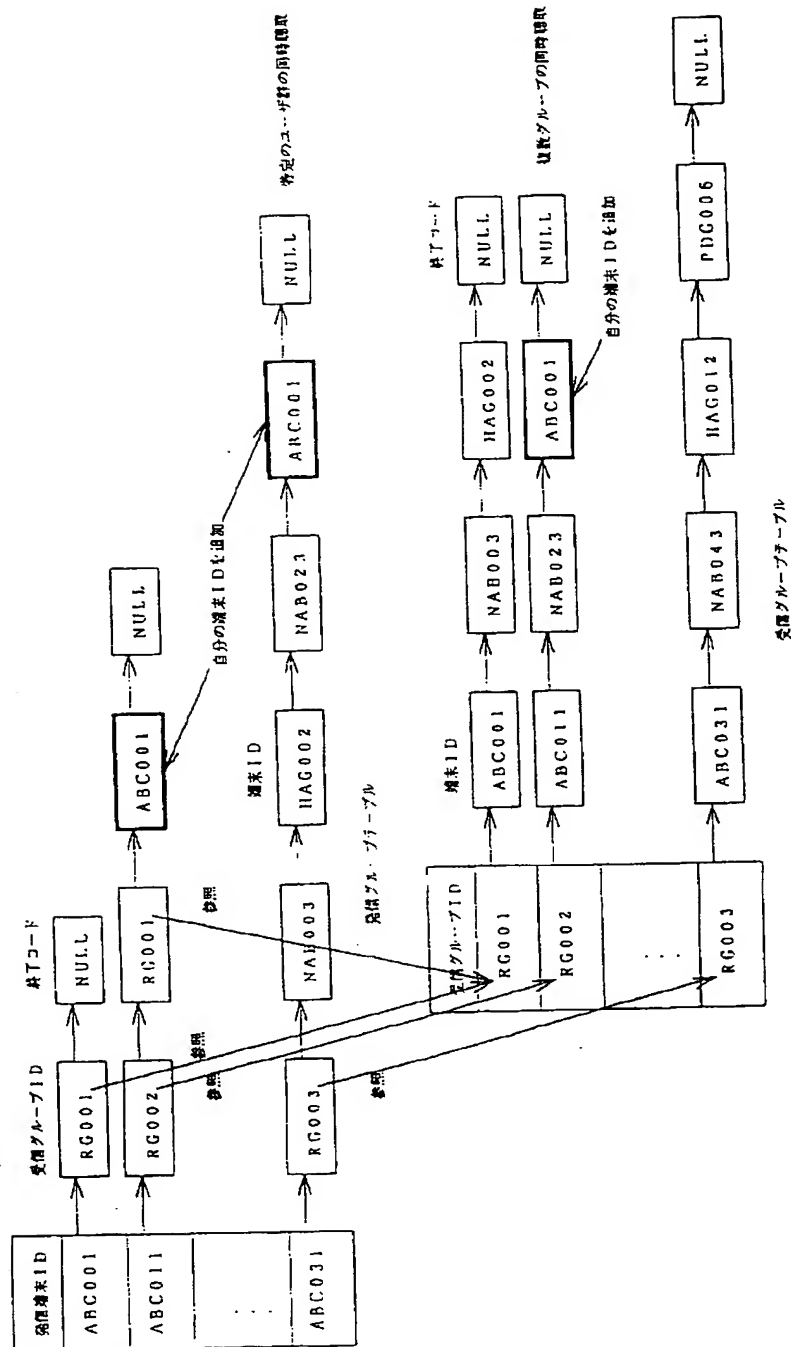
【図43】



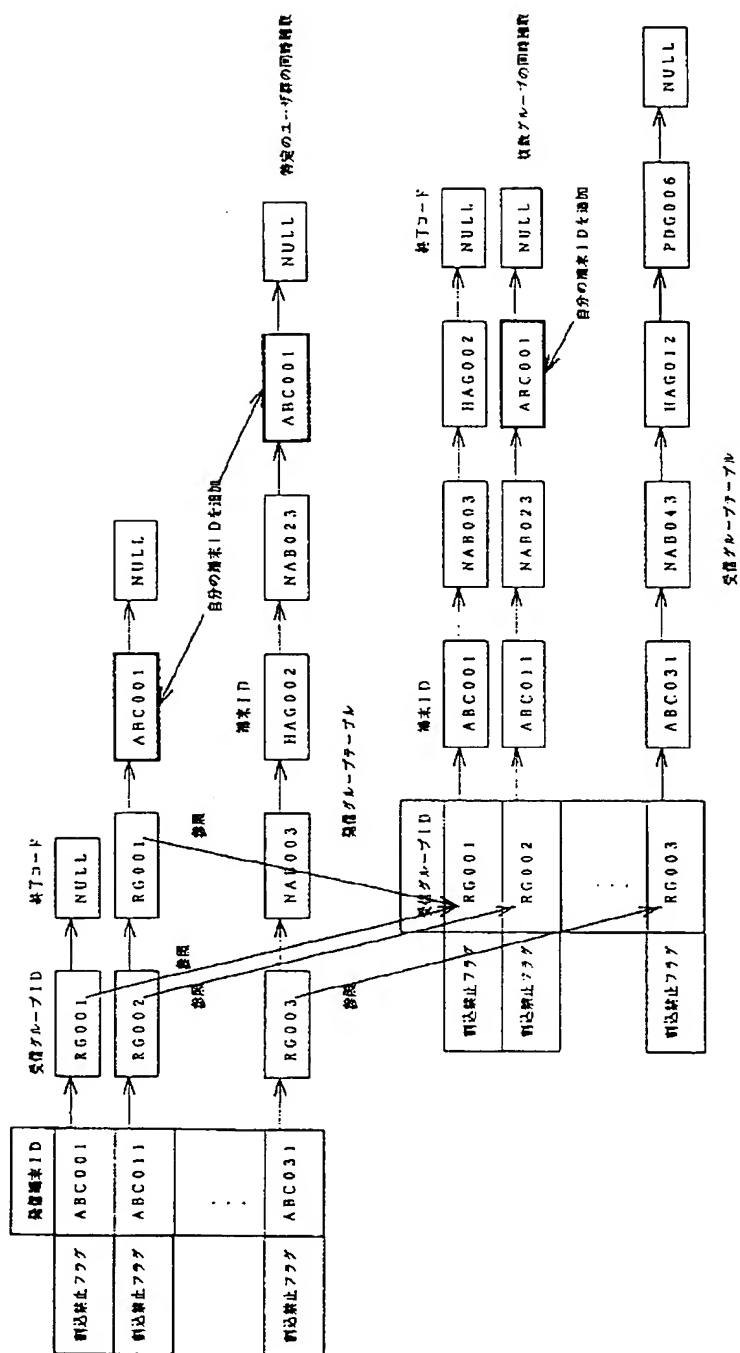
【図40】



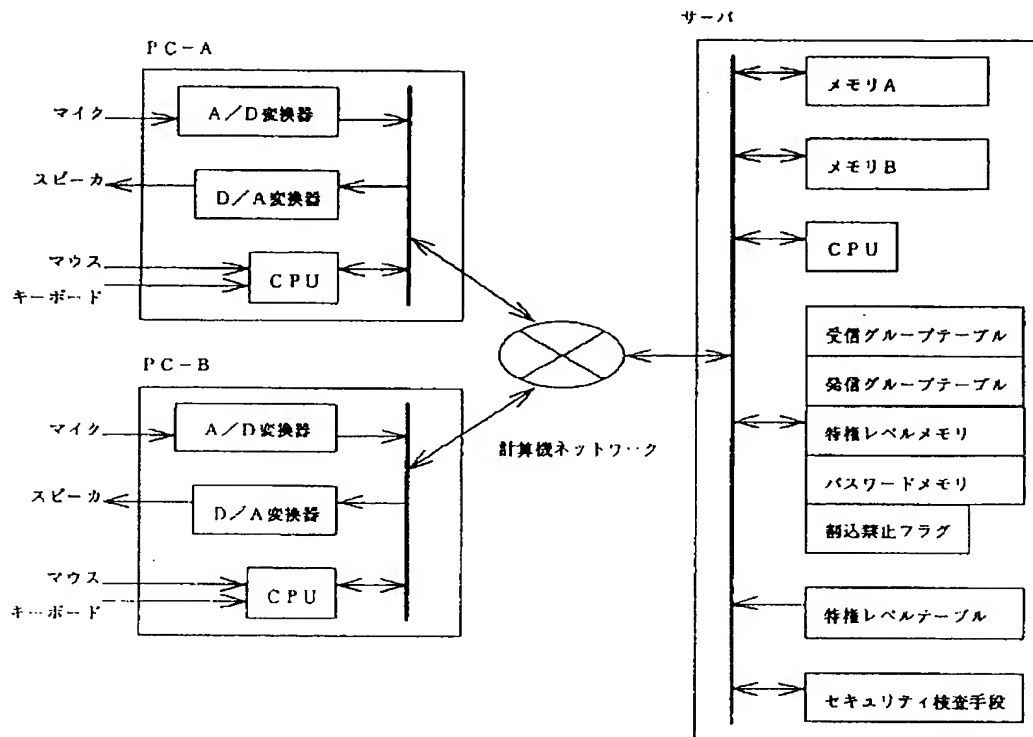
は数ダルの---ブの同時採取



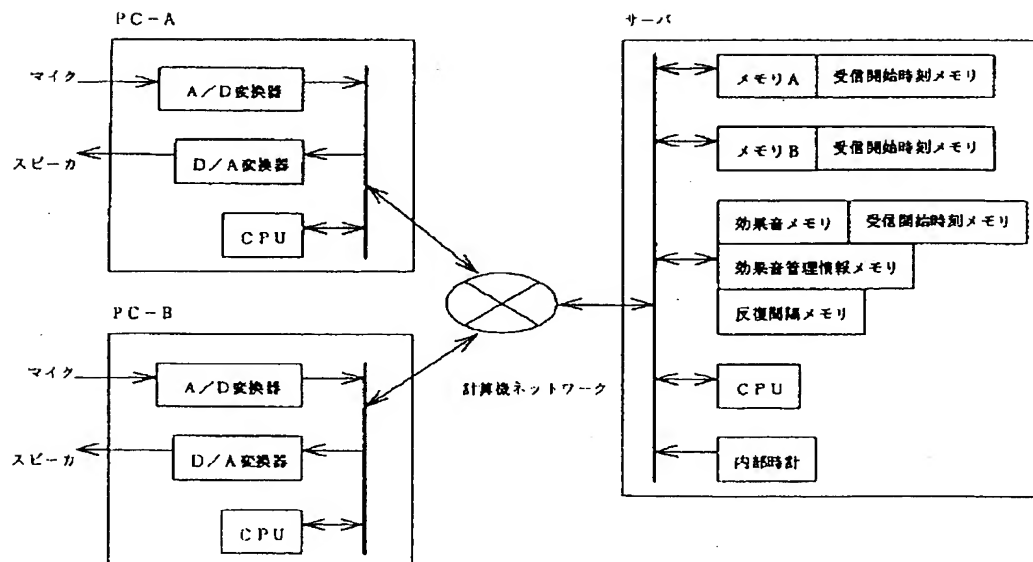
【図44】



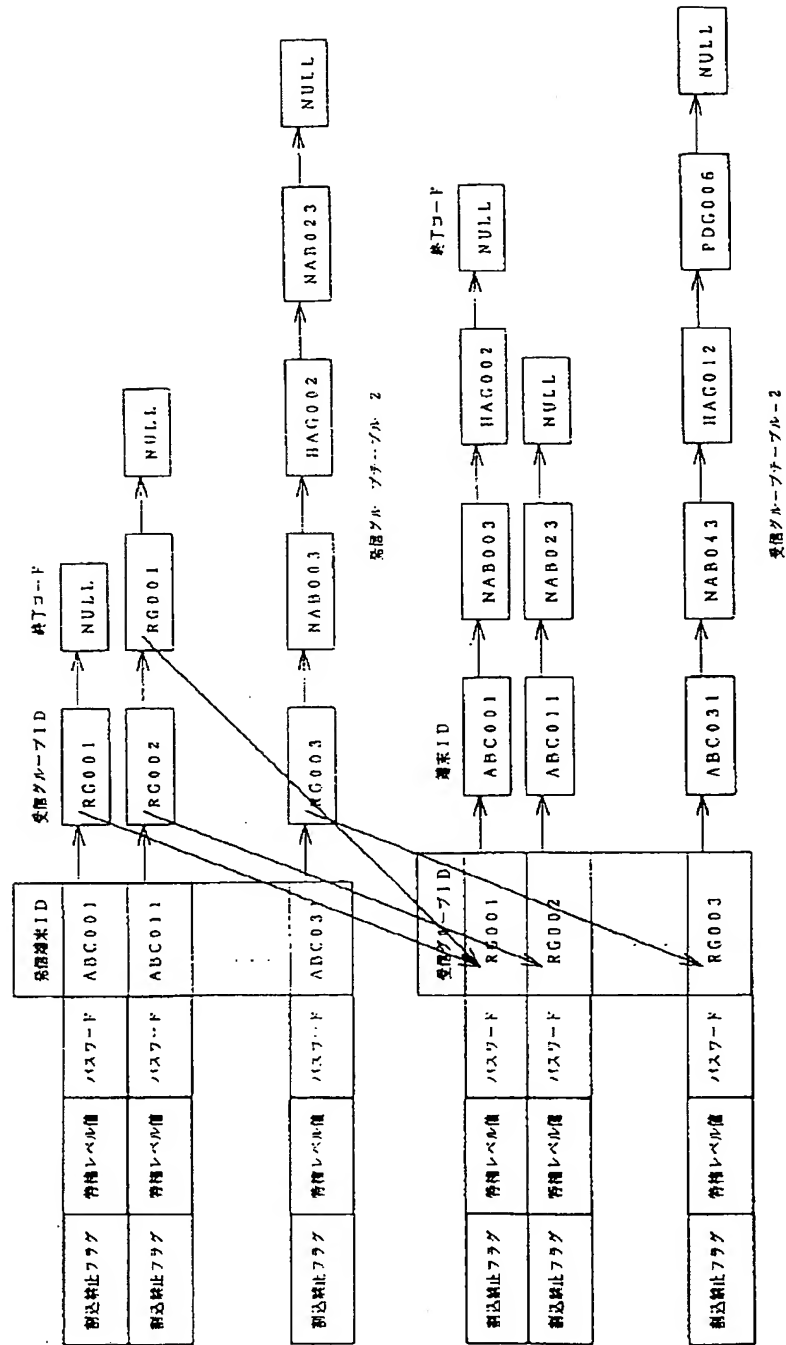
【図45】



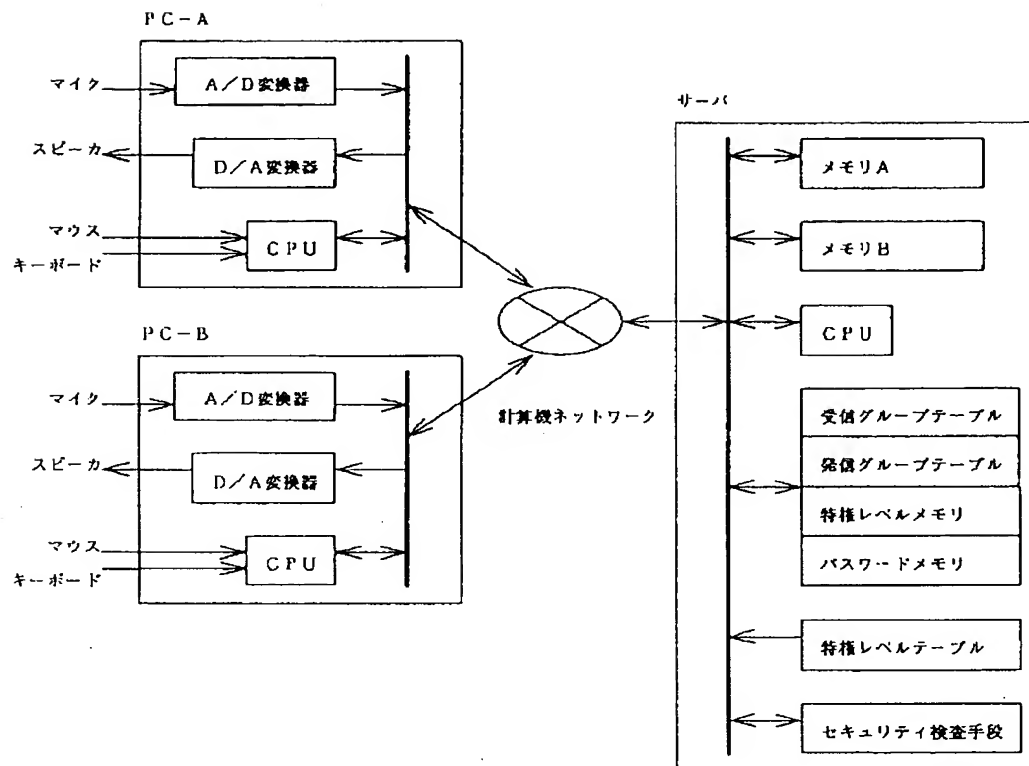
【図52】



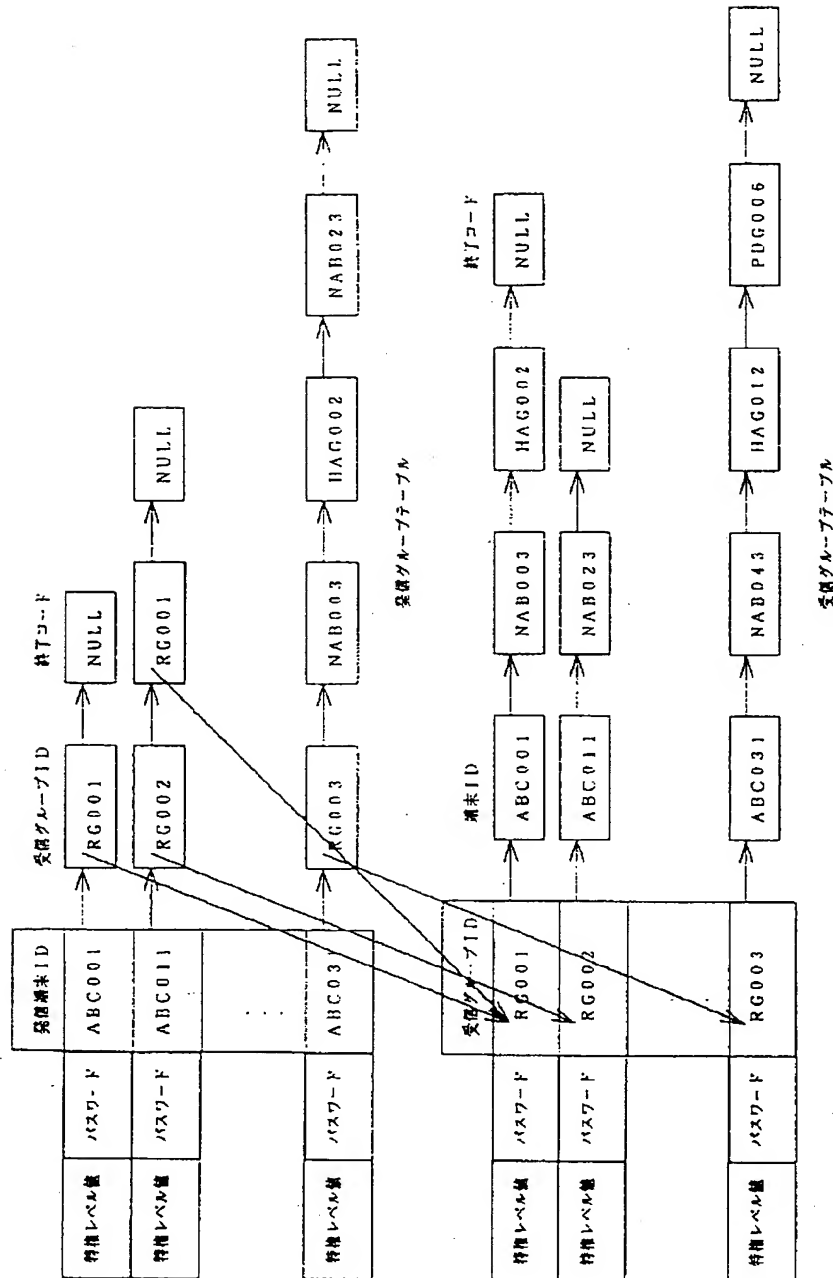
【図46】



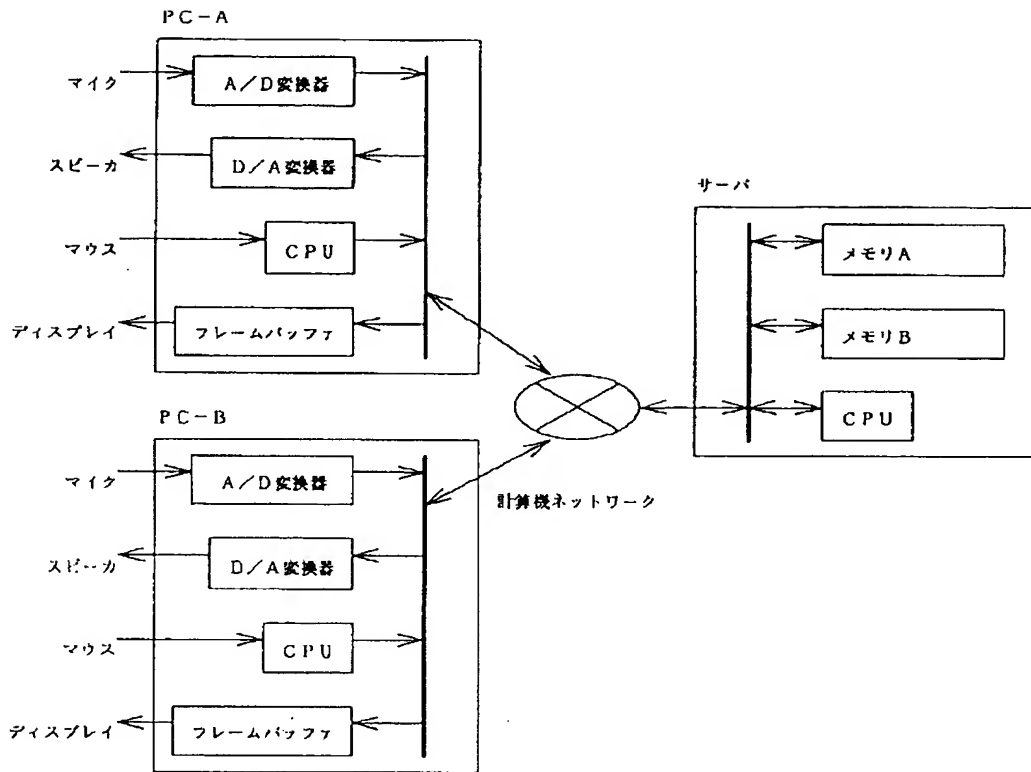
【図47】



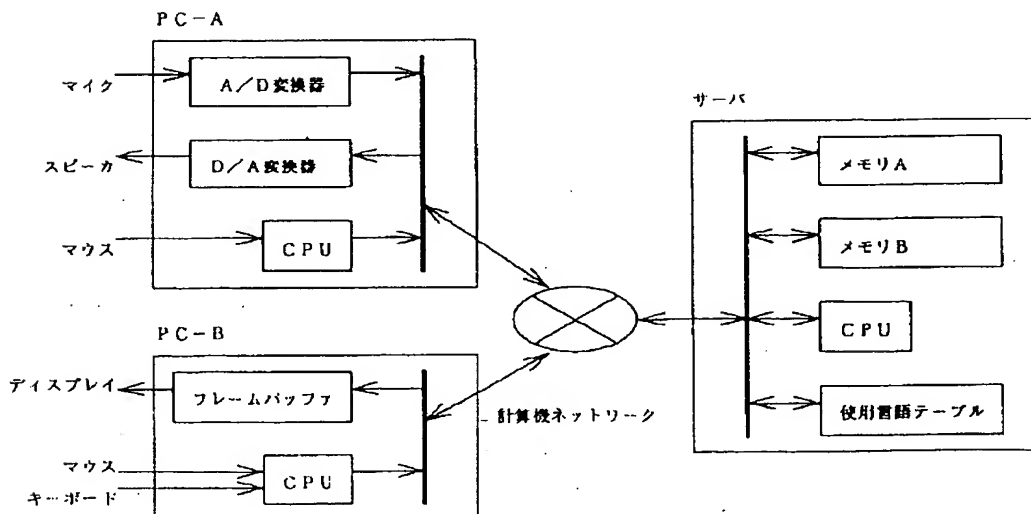
【図48】



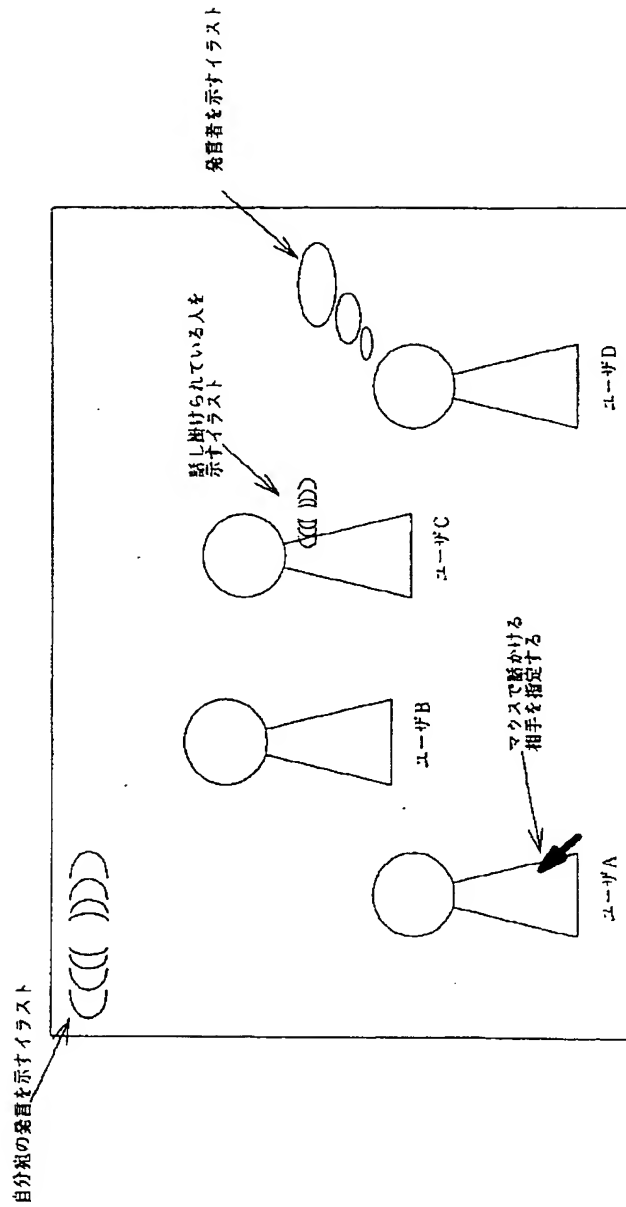
【図49】



【図55】

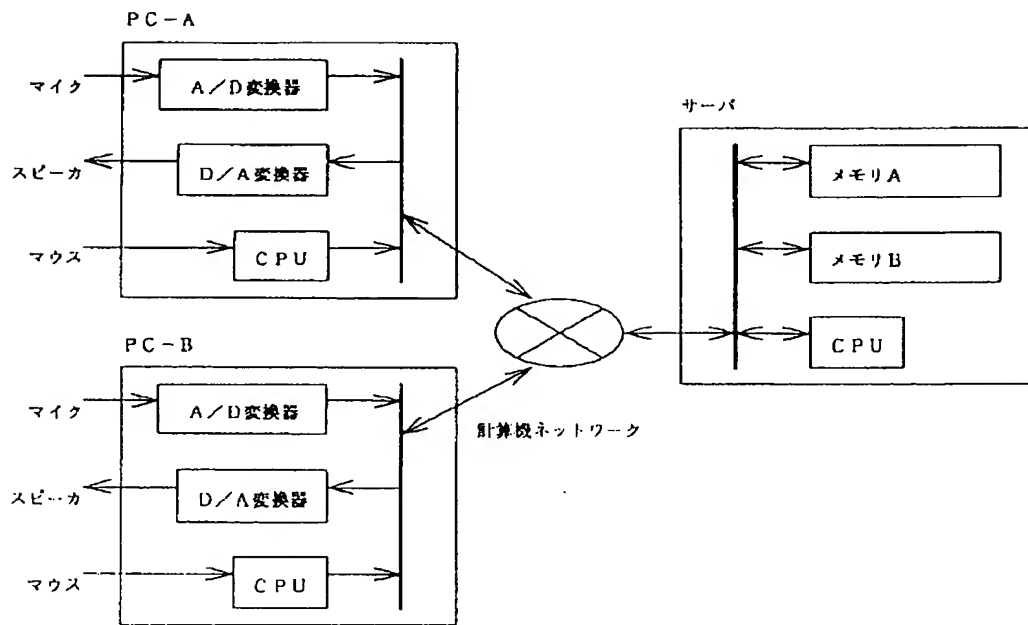


【図50】

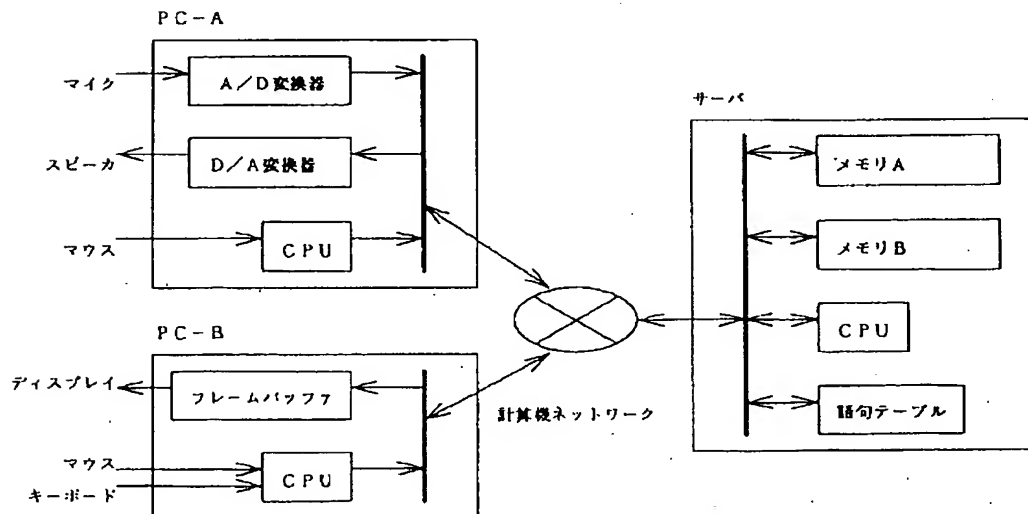


イラストのサンブル

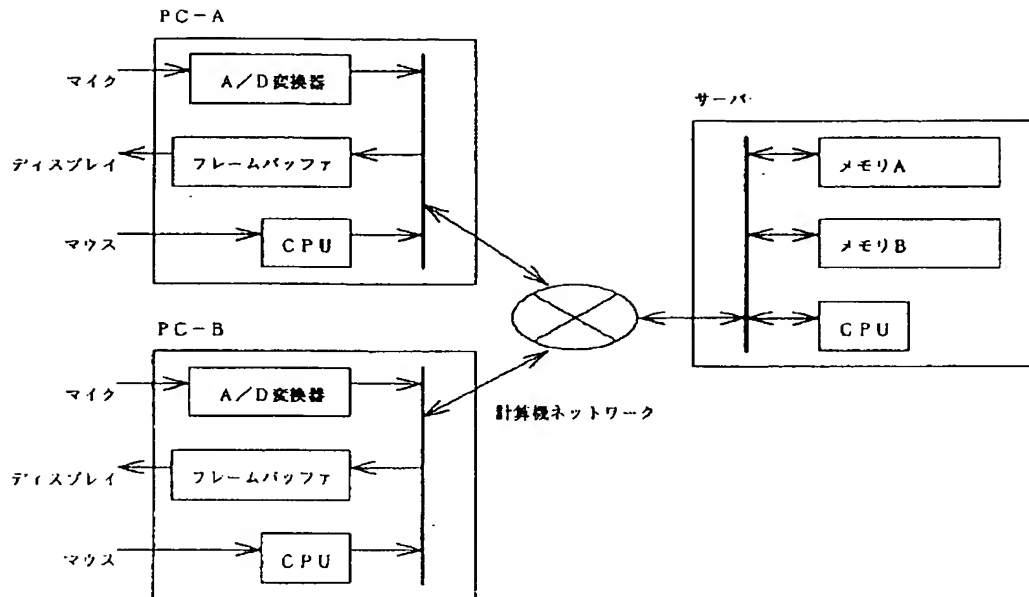
【図51】



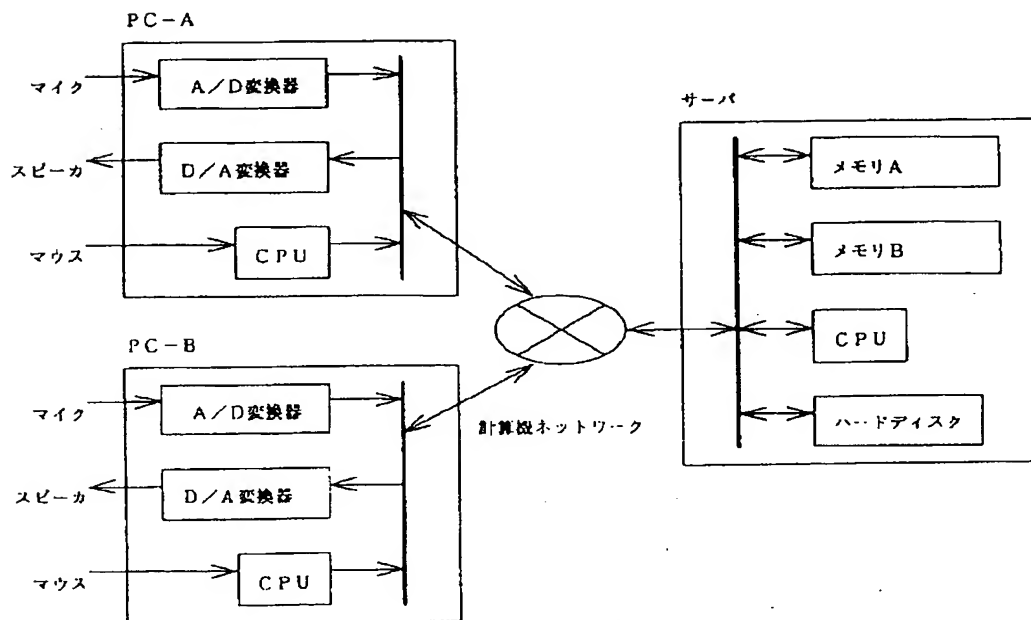
【図56】



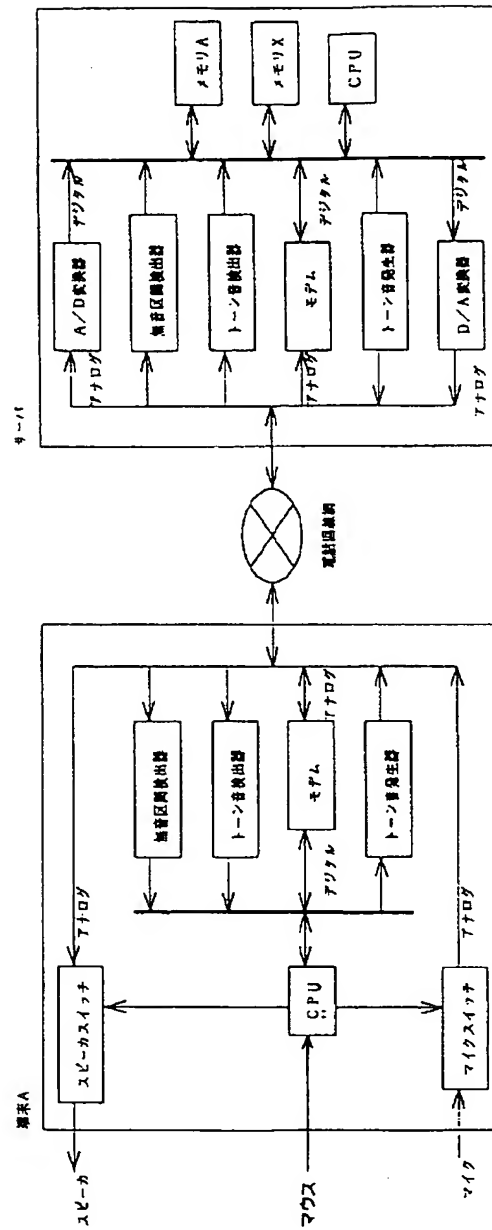
【図53】



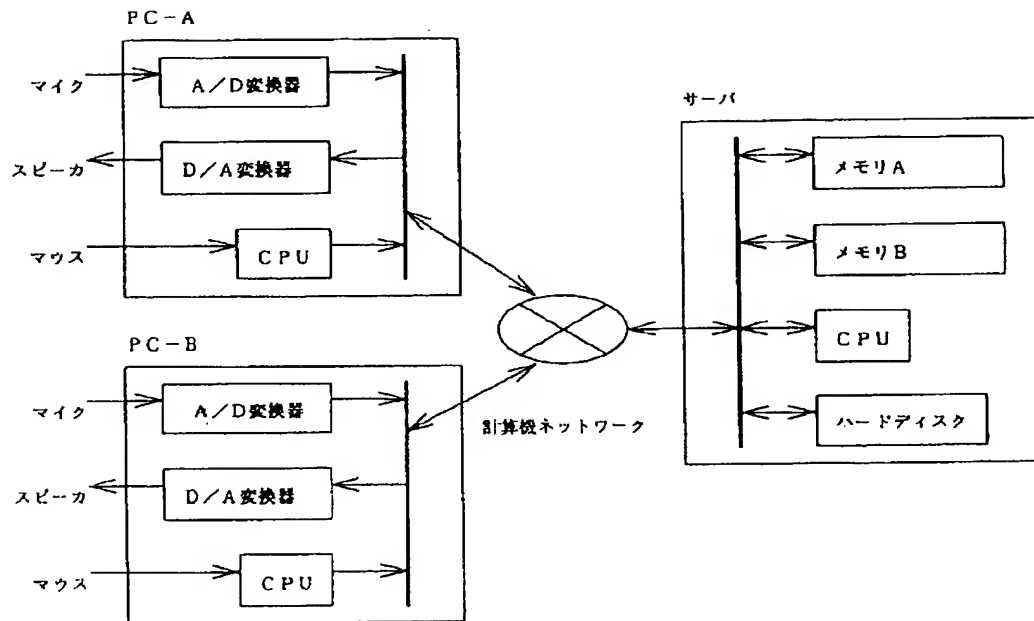
【図58】



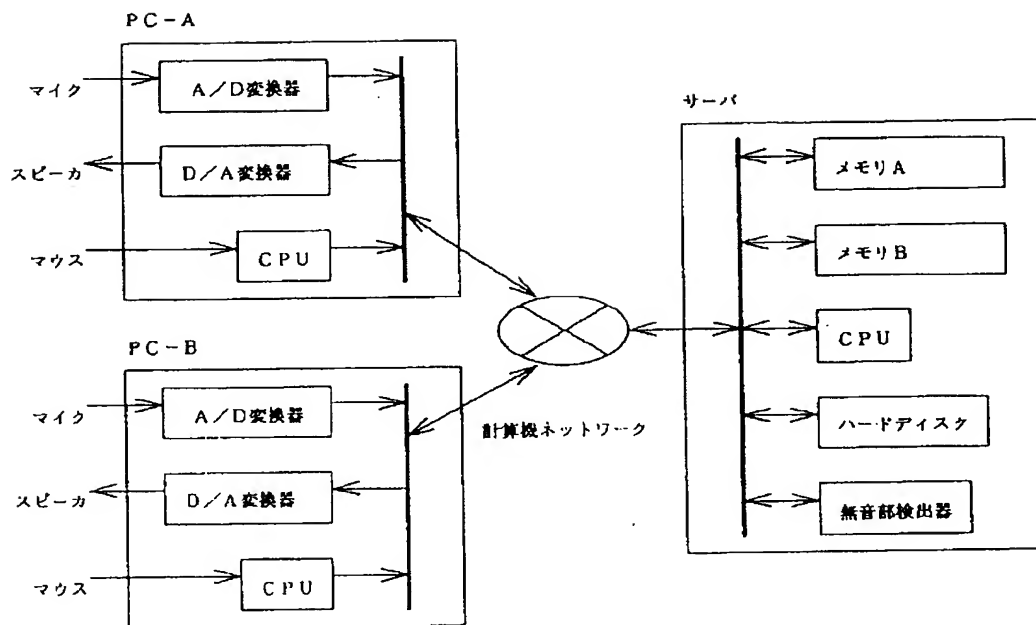
【図57】



【図59】



【図60】



フロントページの続き

(72)発明者 大野 敬史
 神奈川県川崎市中原区上小田中1015番地
 富士通株式会社内

(72)発明者 岩川 明則
 神奈川県川崎市中原区上小田中1015番地
 富士通株式会社内

(72)発明者 原田 裕明
 神奈川県川崎市中原区上小田中1015番地
 富士通株式会社内